EESTI RAHVENÕUKOGU
FISCAL COUNCIL

FISCAL COUNCIL OCCASIONAL PAPER 2

Structural Budget Balance Revisions for Estonia and How to Live with Them (If You Have to)

Ricardo Vicente

March 2022

Eelarvenõukogu
Estonia pst 13, 15095 Tallinn
+372 668 0661
info@eelarvenoukogu.ee
www.eelarvenoukogu.ee

Corresponding author´s e-mail address:
ricardo.vicente@eestipank.ee
ISBN 978-9916-4-1079-0 (pdf)

Structural Budget Balance Revisions for Estonia and How to Live with Them (If You Have to)

Ricardo Vicente

March 2022

Abstract

This paper studies the revisions made by the European Commission, the IMF and the OECD to the data for Estonia on the structural budget balance, the output gap and the budget balance. It shows that the revisions to the structural budget balance and output gap data are large and sign revisions are common, indicating that the structural budget balance data are imprecise and unreliable and the estimates of the output gap are imprecise and of debatable reliability. The paper argues that those estimates are not a good basis for setting and evaluating fiscal policy. Assuming though that they continue to be used, the paper suggests and discusses a simple method for calculating structural budget balance thresholds and a simple algorithm that uses a composite indicator of macroeconomic leading variables to assess how trustworthy the structural budget balance and output gap estimates are, and how likely it is that the sign of the structural budget balance will change in later revisions.

This paper presents the views of the author and does not necessarily represent the views of the Estonian Fiscal Council or Eesti Pank. I am grateful to Lauri Punga for data collection, to François Blondeau for an explanation of a methodological change in estimation of Estonian data and to Lauri Punga and Martti Randveer for comments and suggestions. The usual caveat applies.

Contents

1. Introduction

The structural budget balance is one of the key central macroeconomic and fiscal indicators in the European Union's fiscal framework. The structural balance is calculated as the government balance, which is total general government revenues minus total general government expenses, with the effects of the business cycle on revenues and expenses and the value of one-off and other special fiscal measures removed.

The general rule is that the structural budget balance cannot be in deficit by more than 0.5% of potential output for countries with a public debt ratio above 60% of GDP and it cannot be negative by more than 1% of potential output if the public debt ratio is below 60%. National fiscal rules can be more stringent than EU rules, and the general principle in Estonia, for example, is that the structural budget must be balanced or in surplus[1].

The Covid-19 public health and economic crisis caused the EU fiscal framework rules to be suspended from March 2020, and they will remain so in 2022. In January 2022 there are signs coming from the European Commission and other institutions that the Stability and Growth Pact (SGP) may be subject to extensive reform[2]; the President of France and the Italian Prime Minister support a thorough revision of it (Draghi and Macron, 2021).

In normal times, each EU country has a medium-term objective (MTO), which is the maximum structural budget deficit the country may have. If the structural budget deficit exceeds the MTO, the general rule states that the structural budget should be brought back towards balance at the pace of at least 0.5 percentage point of GDP per year.

Respecting the EU fiscal framework thus fundamentally requires there to be reliable estimates of the structural balance for the current and previous years. The estimates should be not only reliable but also very precise, as the MTOs and the required improvement towards the MTO are set in tenths of a percentage point.

However, these estimates are subject to large revisions. The first objective of this paper is to document the revisions for Estonia by calculating their descriptive statistics with data from the European Commission (EC), the OECD and the IMF. The second objective of the paper is to address the lack of precision and reliability by presenting simple methods for gaining some confidence that an estimated structural deficit will not eventually be revised into a surplus, or a surplus into a deficit. The paper focuses more on EC data than on the other sources because they are more available.

Current year estimates of the level of the structural balance and the change in the level are very relevant for informing policy in the current year, as they may give an earlier indication of deviation from the fiscal rules or from the MTO. Previous year estimates are very relevant for assessing whether the fiscal rules were respected, which is one of the central tasks of the EU's independent fiscal institutions, and how policy should be changed to address any deviations from the rules.

If the estimates from the previous year are not very precise and reliable, it is highly unlikely that those from the current year will be.

The structural budget balance is calculated from the estimates of the output gap, which is the difference between actual and potential output divided by potential output. These estimates are very difficult to arrive at, and this is especially true for more recent years, for more volatile economies like

[1] State Budget Act §6 (1). There may, however, be a structural budget deficit if the sum of the structural budget balances in the years since 2014 is positive, though the deficit cannot be larger than this sum and it cannot be more than 0.5% of GDP [State Budget Act §6 (1^1) and (1^2)]. See the current version of the State Budget Act at: https://www.riigiteataja.ee/en/eli/ee/529122017006/consolide/current.

[2] The SGP is a set of EU norms containing many of the rules of the EU fiscal framework.

the Estonian economy, and in extraordinary circumstances like when an economy is hit by a shock such as the Covid-19 public health and economic crisis.

All these and other difficulties cause large revisions to be made to the estimates of the output gap and so to the estimates of the structural budget balance. Such revisions are large, or at least significant, even though "from an ex post perspective, the [European] Commission considers the structural balance to be in line with the MTO if it is within 0.25pp of GDP of its value" (European Commission, 2019, section 1.3.2)[3].

While later estimates may be less likely to be heavily revised, they are not necessarily more reliable since the difficulties in estimating the output gap and so the structural balance are not only technical but also conceptual[4].

Precise and reliable estimates of the output gap are also important in themselves as they are used for setting the correct stabilisation policy. An inaccurate signal that the economy is overheating might for example trigger unnecessary austerity, causing great economic and social damage.

The other important reason why estimates of the structural budget are revised is that the estimates of the general government budget balance get revised. Even though this public finance quantity is conceptually straightforward and easily observable and measurable, current year estimates are revised because of changes in policy and because of the business cycle, and previous year estimates can even be revised as well. These revisions are usually much smaller, but they are sometimes significant relative to the magnitude of the fiscal targets.

The next part of the paper discusses the sources of revision and the data used, the third part describes the methods used to assess the precision and reliability of the estimates and makes that assessment, the fourth part suggests simple methods for improving certainty about the sign of the estimates, and the last part concludes by discussing the results and making recommendations about the fiscal rules. All the tables are at the end of the paper in Appendix 1, and Appendix 2 summarises an algorithm that can give some confidence about the sign of the structural budget estimates.

2. Sources of revision and data

The structural budget balance (SB) is calculated as

$$SB = GB - CC - OO \quad (1)$$

where *GB* is the general government budget balance, *CC* is the cyclical component and *OO* is one-off and other temporary budgetary measures. The cyclical component is calculated as

$$CC = \varepsilon OG \quad (2)$$

where ε is the budgetary semi-elasticity, which is a measure of how sensitive the general government budget balance is to the output gap, and *OG* is the output gap[5]. It follows that the fundamental ingredients for calculating the structural budget balance are the government budget balance, the

[3] While the limit on structural balance is stricter in Estonian law than in EU law, Estonian law only requires an adjustment to be made if the structural budget deficit is more than 0.5pp away from the limit (State Budget Act §7 (2)). At the same time, one of the criteria in EU law for declaring a deviation from the MTO to be significant is that "the deviation is at least 0.5% of GDP in a single year or at least 0.25% of GDP on average per year in two consecutive years" (European Commission, 2019).

[4] The Fiscal Council Occasional Paper 1 provides an extensive, deep and up-to-date discussion of the problems of calculating the output gap (Vicente, 2020).

[5] For a glossary of these terms see European Commission (n.d.); for a rigorous discussion of the budgetary equations see Mourre et al. (2013) and Mourre et al. (2014). Another commonly used related concept is the cyclically adjusted balance, defined as $CAB = GB - CC$.

sensitivity parameter, the output gap, and the one-off measures. Revisions to the structural budget balance are thus caused by revisions to one or more of these ingredients.

The EC publishes macroeconomic data in spring and autumn and, in some years, in winter, with the cut-off dates for the inclusion of new information in late April, mid-October and mid-February. The IMF publishes its estimates in April and the OECD does so in May or June. As those series are longer, this paper focuses most on the EC data.

The spring data are used here because they have high policy relevance, as they provide an early signal of whether governments respected the fiscal rules in the previous year, and whether they are likely to respect them in the current year[6]. The spring data covering the previous year are used by independent fiscal institutions to assess compliance with the fiscal rules. The spring data for the current year can help governments adjust fiscal policy as needed to comply with the rules.

The data used in the paper start in 2005, which was Estonia's first full year in the EU, and they end in 2019. The data for 2020 are not used because of the especially elevated uncertainty about the economy in 2020, which was a consequence of the Covid-19 economic and public health crisis, the exceptional fiscal policies that were deployed to deal with it, and the suspension of the EU's fiscal rules.

The EC only rarely revises the data about one-off and other temporary measures, and only does so minimally, while the budgetary semi-elasticity is only revised every three years after the first revision, so the only two changes were made in autumn 2014 and spring 2019. The EC budgetary semi-elasticity for Estonia was 0.3 until spring 2014 and 0.443 until autumn 2018, and it has been 0.486 since spring 2019[7]. The descriptive statistics of the revisions of the structural budget estimates barely change at all once the effect of the semi-elasticity revision is removed[8].

A revision of the government balance has a direct one-to-one impact on the estimate of the structural balance, while the impact of a revision of the output gap is deflated by the semi-elasticity. So if the semi-elasticity is, say, 0.486, a revision of one percentage point in the government balance leads to a revision of one percentage point in the structural balance, whereas a revision of the same magnitude in the output gap leads to a revision of only 0.486 percentage point in the structural balance.

However, revisions to the output gap are usually much larger than revisions to the government balance, which are in general small or zero (see below), and so it turns out that the output gap revisions contribute more to the revisions of the structural balance than the government balance revisions do. This means that even after the output gap revisions are deflated by the semi-elasticity, they usually impact the structural balance revisions by more than government balance revisions do. For this reason, this paper pays more attention to the output gap data than to the government balance data as revisions to the output gap data are by far the largest cause of revisions to the structural budget balance data.

[6] Winter data provide even earlier signals but they are released only very rarely.

[7] In the data sample in the paper, the semi-elasticity revisions always result in a decrease in the estimates of the structural balance. This does not always have to be the case, as an upward revision of the semi-elasticity when the output gap is negative increases the estimate of the structural balance. An explanation of the timeline of the revisions of the semi-elasticities can be found in European Commission (2018) Part II, section 2.

[8] This is because in the sample with 15 observations, only two are affected by the semi-elasticity revision and some changes in the absolute values of the SB revisions cancel each other out. However, the semi-elasticity revision in the sample was able to contribute as much as 0.31pp to the downward revision of the structural budget balance. This contribution is significant for policy evaluation. The average of the absolute contributions of the semi-elasticity revision for structural budget revision is 0.17pp, calculated from six observations.

Between spring 2020 and spring 2021, there was an important methodological change in how the EC estimated the trend total factor productivity in Estonia[9]. As a consequence, the estimates of the output gap for the years 2005-2019 made in spring 2020 all underwent downwards revisions, which averaged -1.55 percentage points; the smallest revision from spring 2020 to spring 2021 was of -0.63 for 2016, and the largest was of -2.36 for 2011.

The next section looks more deeply into the data and assesses the precision and reliability of the estimates.

3. Precision and reliability of the estimates

A simple way to assess the **precision** of the estimates is to look at how often and by how much they are revised. Their descriptive statistics will help in that assessment. Large *averages for the absolute value of the revisions* mean the precision of the estimate is low.

Large and small averages of absolute values may conceal changes in sign from one estimate to the next. Sign revisions are more likely when the earlier estimate is closer to zero.

For that reason, it is appropriate when considering the **reliability** of estimates to check the frequency of *sign revisions*. If the structural balance moves from positive to negative or vice-versa from one estimate to another, and if this type of revision is very common, then the estimates of the structural balance may lead to inappropriate policy recommendations being made. A negative estimate of the structural balance might call for fiscal consolidation, but this might be a serious policy mistake if the actual structural balance is positive. A wrongly estimated structural surplus may equally lead to costly economic policy mistakes, and a failure to evaluate whether the structural balance has improved or not may similarly lead to the wrong policy prescriptions.

This means that a source of estimates can be considered reliable if the number of sign revisions is low relative to the number of estimates.

3.1. Simple assessment of data precision and reliability

The data on the government balance, the output gap and the structural budget balance that are published each spring by the EC are shown in Tables 1-3.

If the estimates made in the spring of the current year, which are in reality forecasts, are ignored, the estimates for GB change only a little or not at all from one data vintage to the other. This means that the European Commission's GB data are precise.

The sign of the GB estimates almost never changes from the estimate for the previous year. The exceptions are 2013, for which the sign changed in the spring 2020 vintage from -0.2 to 0.2; 2016, for which the sign changed in spring 2018 from 0.3 to -0.3; and 2019, for which the sign changed in spring 2021 from -0.3 to 0.1. These exceptions occur when the estimates are close to zero and involve relatively small revisions. This means that the European Commission's GB data are reliable.

[9] Since its autumn 2020 forecast, the EC has used a shorter series of the capacity utilisation indicator for Estonia that begins in 2002. The series used previously started in 1995. This means the length of the series is the same as those used for Latvia and Lithuania. See European Fiscal Board (2021).

The estimates of the output gap continually move up and down and frequently change very substantially from one data release to another. This means that the European Commission's OG data are imprecise.

Leaving aside the current year estimates and the data vintage from spring 2021, which was affected by the methodological change, the sign of the OG estimates changes five times, but only for two years. The 2005 OG estimate switches to have a negative sign in spring 2007 and then switches back to positive in spring 2008, after which it stays positive for all the other vintages; and the 2011 OG estimate switches to have a positive sign in spring 2013, then switches back to negative in spring 2015, and back again to positive in spring 2020.

The methodological change led to five more instances of sign change, all of which appeared in spring 2021, as the estimate for 2011 changed for the fourth time to become negative again, and the estimates for 2012-2015 all switched to become negative. This meant that six of the fifteen years saw at least one sign change.

Including the current year estimates produces three more instances of sign change for 2008, 2012 and 2016, and this raises the number of reference years with a sign change to eight.

There are 149 estimates for the 15 reference years and the sixteen data vintages, or 135 without current year estimates and 120 if the spring 2021 data vintage is also excluded. The total number of instances of sign change is small relative to the number of observations, as there are 14 in total, or 11 without current year estimates and five if the spring 2021 vintage is also excluded. Sign change affects eight, seven or two of the fifteen reference years depending whether all the data are used or if the current year estimates and the 2021 vintage are excluded.

These numbers may be small because the OG estimates are relatively far from zero, so even large revisions may not be enough to change the sign.

Relative to the number of observations, the European Commission's OG data can be considered reliable overall. Relative to the number of years affected by sign change, the OG data may or may not be deemed reliable depending on how much attention the current year estimates receive when policy is set, and on whether the spring 2021 methodological change is considered exceptional or not. In any case, whether or not it is exceptional, the change should presumably be interpreted as helping make the estimates more accurate, as far as that is possible given that the OG is a controversial concept, and as such it can only be seen as supporting the argument that the OG data are unreliable. Otherwise there would have been no need to make the change in the first place.

A further argument why the data are unreliable is that the large size of the revisions suggests that more years would have been affected by sign change had the estimates been closer to zero. Understanding whether the economy is above or below its potential matters greatly for policy decisions, and if the estimates are of dubious reliability, then those estimates cannot safely be used for informing policy.

The structural budget balance data move up and down and are subject to significant and large revisions. Given the policy importance of these data and the precision required by the fiscal rules, the European Commission's SB data are imprecise.

Though the ups and downs are not as common as they are in the OG data and the revisions are usually not as large, the European Commission's SB data turn out to be very unreliable. As SB estimates tend to be relatively close to zero and as they are affected by the large changes in the OG data, sign changes become very common. Leaving aside the current year estimates and the spring 2021 vintage, there

are still sign changes in nine of the 15 reference years and 14 of the 84 observations, and there are four reference years in which the sign changes more than once.

The methodological change in spring 2021 creates three more instances of sign change, and this raises the number of reference years affected by sign change to 10.

While it is not surprising that the current year estimates change one year later, the four instances of sign change that are added by including the current year estimates correspond to very significant revisions, one of them by more than 4.0pp. This revision raises the number of years with at least one sign change to 11 of the 15.

The following section evaluates the precision and reliability of the data in greater detail.

3.2. Quantification of data precision and reliability

Current year and previous year estimates are very relevant for policy-making, and a first signal how precise and reliable they are is given by their first revisions. This makes the two subsequent revision categories, which follow the methodology in Darvas (2016, 2019), very relevant as well:

- **revision of the current year estimate one year later**: "[f]or example, the final data point is the difference between the spring 2015 and spring 2014 estimates for 2014. Since, for example, the spring 2014 estimate for 2014 involves forecasts, it is not surprising than [*sic*] revisions are larger than the revisions of the estimates for the previous year";
- **revision of the previous year estimate one year later**: "[f]or example, the final data point is the difference between the estimates for 2013 published in spring 2015 and spring 2014".

These categories can more simply be called the *current level revision* and the *previous level revision*. Current estimates or current year estimates are those published in the spring of the year they are for, and previous estimates or previous year estimates are those published in the spring of the year after the year they are for. A zero revision can be defined as one with an absolute value of 0.01pp or less and the nomenclature for revision sizes is suggested in the box below.

Tables 4-6 present the consecutive revisions of the government balance, the output gap and the structural balance, and each row corresponds to the difference between two consecutive estimates. Current level revisions are shown in row *t + 1*, and previous level revisions in row *t + 2*. Row *t + 3* gives the difference between the estimates published three and two years after a given year, so for, say, 2006 these are the estimates from spring 2009 and spring 2008, while row *t + 4* shows the difference between the estimates published four years after the reference year and those published three years after, so those from 2010 and 2009 for 2006, and so on.

The GB data are fairly precise overall, except for the current level revisions, with many consecutive revisions being essentially zero. The current level GB revisions are large or very large in 2006-2012 and 2014 and are significant in three other cases, but over time the GB revisions seem to converge to zero, even though non-zero revisions are common and, in some cases, significant after the first zero revision. If the current year revisions are excluded, the non-zero revisions are never large.

Unsurprisingly, the previous level revisions are much smaller than the current level revisions, but in some cases they can be significant. The absolute value of revisions from each previous year estimate to the spring 2021 estimate is on average 0.31pp, with a maximum value of 0.91pp for 2006; if the years 2005-2009, which cover a deep crisis, are excluded, the average becomes 0.24pp and the maximum is 0.68pp in 2016.

Comparing the absolute value of the GB current year revisions with the absolute value of the OG current revisions shows that the GB revisions were larger than the OG revisions in seven of the 14 years; there are no figures available for 2005. The GB current year estimates are not much less imprecise than the OG estimates.

It can be seen from the OG and SB revisions that a very large number of revisions are very significant for the reference values of the fiscal rules for the structural budget deficit and for the minimum improvement towards the MTO of 0.5pp of output per year, and that the picture is worse for the OG revisions.

The OG revisions are in general significant, large or very large not only at *t + 1* but across all the time horizons, and they seem never to converge to any value, and definitely not to zero. Indeed the OG revisions are almost never close to zero. The OG estimates for each year change continually up and down over successive spring publications. Putting this evidence together with the discussion about reliability, it is legitimate to say that the OG estimates are very imprecise, their reliability is debatable, and there seems to be no real hope of ever getting an estimate that does not suffer further significant revisions. The impact of the methodological change exacerbates this, but even without it the imprecision remains. More evidence about the OG estimates is presented further on in this paper.

The SB revisions inherit the properties of the OG revisions and so they do not stabilise over time, but move up and down and experience significant revisions even some years after the first estimate, with some of those revisions very significant because of the spring 2021 data vintage. Current year revisions are almost never small, and previous year revisions are small in four out of fifteen cases. The number of consecutive revisions close to zero is zero.

Unlike with the OG data, the case for unreliability is clear cut since sign changes affect nine of the 15 years even without the current year estimates and the 2021 data, with four reference years seeing more than one sign change. If the 2021 vintage is included, the number of years affected increases to 10, and with the current year estimates that number goes up to 11. The sign changes may be more frequent because the estimates are usually closer to zero than the OG estimates are.

Even after the current level revisions and those affected by the spring 2021 methodological change are excluded, there is at least one revision of more than 0.25pp in absolute terms for every year except 2012 and 2014. But even for these two years the estimates are imprecise or even meaningless, as the 2012 SB estimate moves from 0.17 in 2013 to -0.16 in 2020, which is a significant change in absolute terms and also because of the sign change. These numbers are seemingly meaningless though because the latest estimate of 0.92 not only changes the SB sign again, but increases the value by 1.09pp. The 2014 SB estimates in 2015 and 2020 are essentially the same, but this apparent constancy conceals the move in the SB from surplus to deficit and back to surplus within that short period of time. On top of this, there was a revision of 0.69pp from 2020 to 2021.

This means the SB estimates are imprecise, though not as imprecise as the OG estimates, and they are clearly very unreliable. This makes the SB data unsuitable for their central role in the fiscal rules and for informing fiscal policy. More evidence for this is presented further on in this paper.

Table 7 compares the contributions of the OG estimate revisions and the GB estimate revisions to the SB revisions by taking the absolute values of the consecutive OG revisions, deflating them using their semi-elasticities, and then subtracting the absolute consecutive GB revisions.

It turns out that the OG revision mostly makes a larger contribution to the SB revision than the GB revision does. There are a few exceptions though, and most of them correspond to current year revisions (*t + 1*), where the GB contribution is larger than that of the OG in 11 of the 14 cases. For some reason, some of the other exceptions appeared in the spring 2015 data issue.

Tables 8–10 present descriptive statistics for each reference year using the absolute value of consecutive revisions, and calculated over the multiple spring publications but excluding the current year estimates. This is because current year revisions tend to be the largest or among the largest, as tables 4–6 showed, which is quite natural. This is especially the case for the GB data, while the later revisions to the OG data can be very large and can even be larger than some of the very large current year revisions. Tables 8 and 9 also include a lower panel with statistics for the period after but not including the first time a zero revision was found, if that occurred; as there has never been a zero revision for the SB data, table 10 has no lower panel.

The average consecutive GB revisions for each year are always small in absolute terms, except in 2019, and the maximum revision is always either significant or small.

Box. *Suggested nomenclature for the absolute size of the estimate revisions*

Given that the usual minimum improvement towards the MTO is 0.5pp of GDP per year and that the EC applies a margin of error of 0.25, it is logical to define a structural balance revision of 0.50pp, or more conservatively of 0.75pp, as large, and to consider a revision of 0.25pp or less as small. Revisions between 0.25 and 0.75 may be labelled as significant, and anything above one whole percentage point is very large. That gives the following nomenclature for the absolute size of the estimate revisions:

0 to 0.25pp: *small*

0.25 to 0.75pp: *significant*

0.75 to 1: *large*

above 1: *very large*

In stark contrast, the OG revisions are on average never small and large or very large averages are common, with the maximum for each year being a long way above a whole percentage point except in 2015-2017. Of the 15 years considered, 2010 and 2013 are the only ones where the absolute OG revisions manage a zero. The average of the absolute consecutive revisions after the zero is 0.73pp for 2010 from 9 observations. The first and only revision after the zero for 2013 is 1.83pp.

The average SB revisions are not as big as the average OG revisions, except in 2006, and two thirds of them are significant, meaning neither small nor large. There are no zero revisions.

Tables 11–13 keep each time horizon fixed and present descriptive statistics across the reference years, so that row *t + 1* refers to revisions between the reference year and one year later, calculated across different data vintages, row *t + 2* refers to revisions between one and two years after the reference year, and so on.

The average absolute consecutive revisions tend to decrease from the earlier time horizons to the later ones, until they become insignificant for the GB data but not for the OG and SB data. The average of the previous level revisions in row *t + 2* is much lower than that of the current level revisions in row *t + 1* for GB and SB but not for OG. Clearly the SB estimates do not seem to improve as time passes from the first estimates.

In summary, the GB current estimates are not precise and not reliable, but the data deserve trust starting from the previous year estimates. The OG estimates do not ever seem to be precise, their reliability is very debatable and there does not seem to be any perspective from which these estimates

could be deemed trustworthy. The SB estimates are not precise and not reliable for the important policy purposes that they are calculated for.

These conclusions are based on one source, the European Commission, but the other data sources may prove better.

3.3. Comparison across sources

Tables 14–16 present the data available from the OECD and tables 17–19 those from the IMF.

Most revisions of the GB data by the OECD and the IMF other than the current year revisions are zero or small, and in the very few cases where there is a sign change in consecutive estimates, the initial estimate is never greater than 0.40 in absolute terms. The GB data from the OECD and the IMF are precise and reliable, with the exception of the current year estimates.

The OG estimates in both sources are subject to very large revisions, with the IMF revisions generally larger than those of the OECD. Sign changes never happen except in the sole case of the previous year estimate revision for 2017 in both data sources. This means the OG data from the OECD and the IMF are very imprecise, and the IMF is more imprecise than the OECD, but both sources are very reliable in the sense used in this paper.

Tables 20-23 present a range of descriptive statistics for the SB data of the OECD and IMF.

Leaving aside the current level revisions (tables 20 and 21), the SB revisions of the OECD are generally significant, while those of the IMF are generally small or significant. However, the IMF's sample is much smaller and for most reference years the first consecutive revision available is for several years after the reference year. Hence the SB data of the OECD are not precise and those of the IMF are imprecise but less so, but there is only a fairly limited number of observations. The estimates do not seem to stabilise at any point, and later absolute consecutive revisions can easily be significant, large or very large in the OECD data, or significant in the data of the IMF.

Without the current year estimates, there are four instances of SB sign change, affecting four reference years in the OECD data. Previous year estimates never suffer a sign change one or more years later, with the sole exception of the previous year estimate for 2011, which was very close to zero at -0.08, and was sign-revised three years later to 0.10, after which it kept that sign. The OECD SB estimates seem reliable.

The SB previous year estimates by the OECD are in general larger than those of the EC or the IMF, and this could explain why the OECD SB data are reliable. Since it is not possible to say whether the OECD generally overestimates the SB or the EC and the IMF generally underestimate it, then it is not possible either to say whether or not the apparent reliability of the OECD SB is just the result of systematic overestimation. The evidence for overestimation is mixed. The SB data of the OECD are often subject to significant revisions, and such imprecision argues against trusting the signs. However, the OECD revised its data upwards from the previous year estimate to the latest estimate in five of the nine cases, which could imply that the previous year estimates were not overestimated, and that the previous year estimates of the EC and the IMF were underestimated by even more than they seemed to be in comparison with the OECD data.

The reliability of the estimates is very important for assessing the position of the economy and for setting policy but from a technical point of view it is secondary in importance to precision, as there is little value in estimates that never change sign if their levels are far from accurate in the first place.

In the IMF data there is no sign change for any previous year estimate one or more years later, but this assessment comes from only four reference years, 2015–2018. There are only two instances of sign change in the whole dataset, occurring for the SB from 2005 from spring 2017 to spring 2018, and the SB from 2006 from spring 2018 to spring 2019. This suggests the SB data from the IMF are reliable, but this might also just be because of the data limitations.

The SB data of the EC are somewhat more imprecise than the OECD data, and are more unreliable since sign changes happen in 10 of the 15 reference years if current year estimates are excluded.

Overall, the SB data from the IMF seem to be the best, though the sample is much smaller, and so comparisons are difficult. This suggests the EC is the worst data source. For some earlier years, the OECD published their first estimates later and so has access to more information, which might explain why those estimates are slightly superior to the data from the EC. The first estimate for 2005 for example was published in 2012 by the OECD but in 2006 by the EC.

Different sources may not agree on the sign of the estimates. This can be examined by taking the data from the EC and the OECD, as there are more data from them, and looking at the main variable of interest in this paper, the SB, for the years 2005 to 2019, using all the spring estimates published by the EC and the May or June estimates of the OECD published between 2012 and 2021, except 2020 because the OECD did not publish estimates then. This gives the two samples exactly the same reference year-publication year combinations, and with the current year estimates kept in, the average estimate of the structural balance from 74 observations is -0.51% from the EC and 0.36% from the OECD.

Generally speaking, the EC sees deficits while the OECD sees surpluses. Both are referring to the same reality with those figures though.

This large difference in the average estimates quite often implies that the assessments of the structural budget balance are sharply different, and so the policy recommendations based on them are quite different. The EC and the OECD disagreed about whether the structural budget was in deficit or in surplus on 27 occasions out of 74 estimates, or 36% of the cases. This is mostly because the OECD estimates are generally higher than those of the EC, and they could even be thought of as the EC estimates plus a constant plus an error, as the correlation between the two datasets is 91%.

In 26 of the 27 cases of sign discrepancy, the EC finds a deficit while the OECD finds a surplus, the exception being the spring 2015 estimates of the 2010 structural balance.

Sign disagreement in some data vintage or vintages affects 10 of the 15 reference years.

In five consecutive reference years, 2011-2015, out of the nine when the EC sees a deficit at some vintage while the OECD declares a surplus, the EC eventually changes its mind and estimates a structural surplus that it then maintains. There are only two observations for the reference year 2005 from the EC, both of them negative, in spring 2013 and spring 2014, while the OECD saw surpluses. However, the OECD switched and found deficits in 2016 and 2017, before reverting back to surpluses in 2018-2019, and then back again to a very large deficit in May 2021. It is possible that for 2016, there has not yet been enough time for the EC to change its position. The years 2017 and 2018 are interesting because it is the OECD that changes its mind and confirms the deficits first registered by the EC, having disagreed in the first vintage.

There are thus signs that the EC is more pessimistic in its earlier estimates than the OECD is.

Changing position can take time though, as it took seven years from the previous year estimate for the EC to switch from deficit to surplus for 2013 for example, while the OECD was claiming a structural

surplus in all its spring publications, taking the *reference year-publication year* combinations for which there are data in both sources.

The case of 2010 is interesting too, since both institutions initially estimated a deficit and then both changed to surplus and kept that assessment up to spring 2021, with the switch by the EC coming one year ahead of that by the OECD.

Since the correlation between the estimates of the EC and OECD is 91%, all this evidence put together suggests that the differences between the EC and the OECD are the result of the OECD estimates being on average larger than those by the EC.

That the EC publishes its estimates earlier in spring and the OECD later might give some advantage to the OECD, as their information set for the current year is a bit larger.

In any case, it is not clear whether the EC underestimates the structural balance or the OECD overestimates it or both. What is clear is that all these differences mean that neither source can be used alone to lend greater credence to the estimates of the other source. This means that policy makers would do better to look at a different, better and richer source of data such as a macroeconomic heat map (discussed later in this paper).

The differences between the EC and the OECD in their estimates of structural balance are for the most part the result of differences in estimates of the output gap, which can be very large, while the GB estimates are essentially identical for the two sources, except for current year estimates.

The OG estimates might differ because different methods are used for the same output gap concept, or because different institutions have different concepts of the output gap. This discussion, however, only shifts the difficulty in assessing the precision and reliability of the estimates from a comparison of institutions to a comparison of concepts, which is even more difficult to make because it is extremely difficult to tell which concept, if any, is more correct, given that the output gap is essentially a conceptual quantity in any case[10].

Tables 22-23 are comparable to table 13. Once more, the SB data from the IMF compare favourably with the OECD data in their precision but this might be because there are fewer observations for the IMF and the first estimate for most reference years was available at a later date; the first estimate for 2005 for example was only published by the IMF in 2016 while the OECD had already published its estimate in 2012. From tables 22 and 23 it is not clear which if either of the IMF and the OECD is less imprecise.

Tables 24–26 summarise a comparison across the three data sources using the absolute values of the estimate revisions and focusing on the current year estimates and previous year estimates, both of which are central to the work of policymakers and independent fiscal institutions. Two other revision categories are introduced in the tables, following Darvas (2016, 2019):

- **revision of the change from the previous year to the current estimate one year later**: "[f]or example, the final data point is the difference between spring 2015 and spring 2014 estimates of the change in the structural balance from 2013 to 2014";

- **revision of the change from two years before to the previous year one year later**: for example, the final data point is the difference between the spring 2015 and spring 2014 estimates of the change in the structural balance from 2012 to 2013[11];

[10] See the reference in footnote 4.
[11] This revision category is not calculated in Darvas (2016, 2019) but follows naturally from his work.

These categories can be called *current change revision* and *previous change revision*. For the previous and current change revisions, a sign revision to positive means that the later estimate says the structural balance improved where the earlier estimate had said it had deteriorated, and a sign revision to negative means it was later said to have deteriorated having been earlier thought to have improved. For both level revisions and change revisions, a large number of sign changes to positive may be a signal of initial forecasting pessimism, while revisions to negative sign indicate initial optimism.

Across the three variables, the four revision categories and the three data sources, the only small averages for the absolute revisions of the estimates are for the previous level and previous change revisions by the EC and the OECD of GB, which is the easiest variable to estimate. All the other averages range from significant to very large. Some of the maximums are colossal.

At these revision time horizons, no institution seems to be clearly more precise than any other with the GB data; it might be that the least imprecise OG and SB data are those of the IMF and the most imprecise are those of the EC. The averages are generally larger in the EC data, but this is the only source that includes 2005–2007, which were the years with the largest average revisions, together with 2008 and 2012, and the largest revision variation in the OG data, which is the main factor in the revision of SB data. This makes it difficult to compare the EC data and the other two data sources. The IMF samples are shorter, and that makes comparative assessment of that source difficult.

Not surprisingly, the estimates for current level and current change are clearly worse in their precision than the previous year estimates. With three exceptions, the average absolute revisions of the output gap are larger than those of the government balance.

All data sources report sign revisions and no source seems clearly more reliable than the others in these revision categories; the EC reports more sign revisions but its samples are longer.

In general, sign revisions are more common for the GB data than for the OG and SB estimates, as the GB estimates are in general close to zero and as the government might take the current year estimates of GB more seriously than the current year estimates of the OG and the SB, and might react more strongly to them.

There are more sign changes with the SB data than the OG data. In general, the current level and current change estimates are sign-revised more often because they involve forecasts.

For the previous level and previous change estimates for the SB, revisions are in general more frequently made to the negative than to the positive. The previous level estimates from the OECD and the IMF seem reliable as there are no sign changes, but the same cannot be said about the previous change estimates. In any case, the OECD and the IMF samples are short and do not include the earlier years for which estimation might have been particularly difficult. The estimates of the level and change of the SB for previous and current years by the EC are not reliable. No institution is reliable for the current level and current change estimates of the SB, which should not be expected in any case given that such estimates depend on forecasts.

In short, the estimates are not precise enough for evaluating current and previous fiscal policies, for assessing their compliance with fiscal rules, for setting the necessary adjustments towards compliance with fiscal objectives, or for deploying the correct macroeconomic stabilisation policy.

4. How can the signs of estimates possibly be trusted

4.1. Confidence Thresholds

The lack of precision and reliability in the estimates of the structural budget balance poses the question of what confidence thresholds could be set so that if an estimate is higher than the upper threshold or lower than the lower one there would be reasonable confidence that the structural balance is indeed in surplus or deficit. This section looks at EC data covering the years 2005-2019 and published in spring 2006 to spring 2021 to answer this question. The notation used is:

$SB_{t,\,t+n}$ is the estimate of structural balance for year t published in the spring of year $t + n$.

One way to suggest these thresholds might be to look at descriptive statistics for revisions of estimates conditional on some early indication that the structural budget balance is in surplus or deficit. Tables 27-30 present descriptive statistics for the revisions of the current year and previous year estimates one, two and up to 11 years later, conditional on the current year or previous year estimate being negative or positive.

Unlike in tables 11-13, the descriptive statistics in tables 27-30 are calculated over the revisions rather than the absolute value of the revisions, and they are not consecutive revisions but rather they keep using the same *initial estimate* for the current year or the previous year. In table 27 for example, which shows current year estimates, the statistics in the $n = 5$ row were calculated using the differences $SB_{2005,\,2010} - SB_{2005,\,2005}$, $SB_{2006,\,2011} - SB_{2006,\,2006}$, and so on. Here, $SB_{2005,\,2005}$ and $SB_{2006,\,2006}$ are the initial estimates. Similarly, table 29 uses previous year estimates and the row $n = 5$ is calculated from the differences $SB_{2005,\,2011} - SB_{2005,\,2006}$, $SB_{2006,\,2012} - SB_{2006,\,2007}$, and so on, where the initial estimates are $SB_{2005,\,2006}$ and $SB_{2006,\,2007}$.

For the case where a current year estimate is negative, shown in table 27, a simple lower threshold for a given time horizon is the negative of the sum of mean and standard deviation. If $SB_{t,\,t} < 0$, we may be confident that $SB_{t,\,t+1}$ will be negative as well, as long as $SB_{t,\,t} < -1.12$, which is the negative of the sum of mean and standard deviation from the *1 year later* row. In the same way, there may be reason to believe that a current year structural balance that is estimated to be in deficit will not be revised into positive territory after two if it is *-1.04* or less or after five years if it is *-1.55* or less.

The implication from this exercise is that even a seemingly very low structural balance of, say, -1.25% of potential GDP might turn into a surplus after perhaps five years even if it is below the thresholds for one year later and two years later. This can happen if the revision five years later is about one standard deviation above the average revision made five years later. The year 2010 is a case in point (see Table 3), as the current year estimate of -2.10 is the lowest in the series of vintages but it was nevertheless revised into a surplus of 0.32 five years later.

There are also the cases shown in table 3 of the large current year deficits in 2011 and 2012 and the not so large ones in 2014 and 2015 that change sign two or more times over different spring publications.

This evidence suggests that a rule of thumb would be to pick a lower threshold, taking *the negative of the largest mean plus one standard deviation across the different time horizons*. The rule of thumb for current year deficits would then be that if $SB_{t,\,t} \leq -3.40\%$ then it is likely that SB_t is indeed in deficit. Note that *3.40* is slightly higher than the largest of the values in the column Mean + S.D. in table 27.

There are caveats to this rule of thumb, firstly that it is not based on theory. A theory to justify it would require at a minimum the assumption that the estimation processes do not change over time, which is not the case. The calculation of the descriptive statistics for the same time horizon implies that data

from different vintages must be used, meaning that not all the observations come from the same data generating process, so theories about the sample mean do not apply. At best we may assume that the estimation processes remain largely the same even though there are methodological changes. We may also hope that later vintages correspond to better data and so the samples being larger should mean that the descriptive statistics are to some degree better.

The second caveat is that the longer the time horizon is, the fewer observations there are in the calculation of some of the descriptive statistics on which the threshold is based. The descriptive statistics for 11 years later for example are calculated from one single observation because of the lack of data, but at most they could be calculated from six observations for the years 2005–2010; they cannot be calculated for any year from 2011 onwards because not enough time has passed since 2021. Being robust across a lot of time horizons implies that very small samples must be used at the longer time horizons. These very small samples might produce thresholds that are too large in magnitude, making the rule of thumb less useful.

In the previous example, the -3.40 threshold is based on the statistics for 11 years later calculated from one single observation relative to 2010. That exceptional 3.38pp revision is not averaged out as there are no other observations for the 11 years later case.

The corresponding threshold ends up not being much use because there are no current year estimates lower than it, but it is not very meaningful to say that we cannot be sure whether a deficit remains if the current estimate is at any value above -3.40, even one as low as -3.0%.

To set an arbitrary limit, we may look at as many time horizons as possible conditional on there being at least four observations from which to calculate descriptive statistics. The rule of thumb then becomes that if $SB_{t,\,t} \leq -1.90\%$ then it is likely that SB_t is indeed in deficit. The current year estimate is lower than this level only for 2010, but the 2010 SB eventually becomes positive.

Finding useful thresholds is difficult because the data are so volatile, especially when the starting point is itself a forecast, as it is with current year estimates. The $SB_{2009,\,2009}$ is -1.00 for example, which is not so low, and it remains negative in all consecutive vintages, while the $SB_{2010,\,2010}$ is much lower at -2.10 but turns positive six years later and remains so in all subsequent vintages.

Given all the evidence above and the variability both small and large in signs that affects the current year SB estimates, it seems wise not to offer any thresholds for current year SBs that are in deficit. Estimates of the structural budget balance are just too unreliable and it might be that no rule of thumb can mitigate that or help reduce uncertainty.

The same caveats and trade-offs apply to the rule of thumb for the other cases, such as positive estimates of the current year SB. The rule of thumb suggested by table 28 and by adapting the previous logic to the surplus case is *the negative of the smallest mean minus one standard deviation across the different time horizons*. In practice, if $SB_{t,\,t} \geq 5.15\%$ *then it is likely that* SB_t *is indeed in surplus*. Again this threshold does not seem to be much use as this value seems extreme and there are no current year surpluses that high, given the highest is 3.50 in 2007, and it is not meaningful to say that we cannot be sure a surplus of 4.00 will remain a surplus. Applying the condition that descriptive statistics should be based on at least four observations does not help, as the threshold would essentially be the same.

As before, it does not seem wise to suggest any threshold for the case of a current year SB surplus.

There may be different ways to suggest meaningful and useful thresholds, but if there are such ways, they do not seem evident.

An alternative rule would be to use the mean and, say, half of the standard deviation, or to use only the mean. These rules are not fundamentally different from one another but correspond to different points along the trade-off between reliability and usefulness. The smaller the absolute value of the threshold is, the more useful it is, but the less reliable. Given the availability of the data, it is possible that an alternative rule would give the same or very similar thresholds; that is the case for the Estonian data from the EC.

It could be suggested that rather than restricting the time horizons because of the very small samples, it might be possible to restrict the descriptive statistics to a period when there was less data volatility. However, it does not appear particularly meaningful to find rules of thumb from normal data that apply only to times when the data are normal.

Previous year estimates are subject to smaller revisions, so some threshold rule of thumb might be more useful with them.

Applying the rule of thumb of *the negative of the largest mean plus one standard deviation across the different time horizons*, to the case of a previous year deficit gives if $SB_{t-1,\,t} \leq -1.70\%$ then it is likely that SB_{t-1} is indeed in deficit (table 29). Not surprisingly, this rule of thumb is less restrictive than the one for current year deficits, but it also relies on one single observation relative to the same year of 2010. The value is low but perhaps not so low that it is not useful at all.

Finding this number needs the data from all years and all publications to be considered but as it is directly based on the revision of only one year (row *n = 10*) it is not a circular process to check the quality of the rule of thumb using the same sample from which it was taken.

The $SB_{2008,\,2009}$ is *-4.10*, which is lower than the threshold, and the 2008 SB never turns into a surplus and all the successive estimates remain close to that initial previous year estimate. This evidence supports the argument that this threshold is reliable.

The previous year estimates of -2.17 for 2018 and -2.40 for 2019 are also below the threshold and never turn into surpluses, but their latest estimates are 0.57pp and 1.12pp higher than the initial previous year estimates, and at -1.61 and -1.28 are above the threshold. It could then be said that there has not yet been a sufficiently large number of successive estimates that are all negative to make a convincing case that the deficit will remain a deficit. The latest estimates being above the threshold could also be a sign that revisions to positive might eventually happen. These pieces of evidence weaken the rule of thumb even though they do not contradict its reliability.

In summary, there is only a little evidence that this rule of thumb might work, and while there is currently no evidence that it fails, there are some signs that it could.

If *the negative of the smallest mean minus one standard deviation across the different time horizons* is used, the rule of thumb for the estimates that find a surplus for the previous year SB is if $SB_{t-1,\,t} \geq 4.00\%$ *then it is likely that* SB_{t-1} *is indeed in surplus* (see table 30). Again, this is an extreme threshold as no previous year SB has ever been estimated at this level. The $SB_{2006,\,2007}$ of *3.30* is the highest previous year estimate. However, the $SB_{2006,\,2021}$ is negative at -1.32, which indicates that the apparently very large magnitude of this threshold might not actually be that large and the threshold might actually be useful. Another point that supports this view is that four of the seven positive previous year estimates, all of which are below the threshold, are negative in the latest data publication, and this occurs despite the methodological change contributing to upward revisions of the SB by strongly revising the OG estimates downwards.

Putting everything together, the thresholds of *-3.40* for current year deficits and *5.15* for current year surpluses seem extreme and not very useful, as they can at best be used as extreme indicators, meaning that if an estimate respects the threshold, we can at least be confident the sign of the SB estimate will not change. The threshold of *-1.70* for previous year deficits might actually be useful and reliable and the threshold of *4.00*, though apparently very large, might actually be useful too. There are caveats in all cases.

A similar rule based only on the means gives the same threshold of *-3.40* for current year deficits, a very similar one of *5.00* for current year surpluses, the same one of *-1.70* for previous year deficits, and a threshold of *2.20* for previous year surpluses.

The descriptive statistics from SB estimates alone can only be used to a limited degree. A more effective solution is the heat map.

4.2. Using the heat map data to increase trust in the estimates

The heat map for Estonia has been published by the Estonian Fiscal Council since 2018 in order to present a broad picture of the business cycle. It is based on a composite indicator (CI) that takes the simple average of a set of standardised macroeconomic variables that are all conceptually simple, accessible and reliable. The variables are core inflation, the average wage, the unemployment rate, the employment rate, vacancies, capacity utilisation, demand in construction, demand in industry, demand in services, and economic sentiment (Fiscal Council, 2018).

If the messages from the CI and the OG are not consistent then it may be more likely that the OG will be revised until it gives the same message as the CI, implying that the SB will also be revised. There is some evidence that is the case (shown below). If the CI suggests for example that the economy is overheating while the OG says the economy is operating at its potential, then an upward revision of the OG may be more likely, forcing a downward revision of the SB. This can easily explain why an SB estimate turns from surplus into deficit and can help predict that event.

The CI is based on standardised variables that use long-term averages calculated over quarterly data starting in 2007Q1. Vintages of the CI can be calculated, so the 2010 vintage for example can be calculated with data for the period 2007Q1-2009Q4 that was available in 2010. The CI is made annual for each vintage by taking the average of the quarterly CIs of the same year. A value of, say, 1 or -1 means that the CI is one standard deviation above or below its average in a given year, which indicates that the economy is above or below its potential.

Data are available for the period 2007Q1-2021Q1. As before, I exclude the years 2020 and 2021 because of their very exceptional character. The 2008 vintage is not presented because, by construction, the CI for 2007 must be 0.0; in a similar way, the 2009 vintage is not presented because the construction of the CI requires an average of zero over the years, so if there are only two reference years, 2007 and 2008, then one must necessarily be positive and one negative. This meant the earliest vintage I could use was 2010, and to improve a little on that, I start from the 2011 vintage, which is based on data from four years. As a consequence, I am considering for the previous year estimates the reference years 2010–2019, which was a period of recovery from the deep recession in 2008 and 2009; real growth was -5.1 in 2008 and -14.4 in 2009, and it was positive in 2010-2019 (source: IMF, 2020). The CI vintages are presented in table 31.

The notation below is used in this subsection:

$OG_{t,\,t+n}$ is the estimate for the output gap in year t published in the spring of year $t + n$ and

$CI_{t, t+n}$ is the composite indicator for year t calculated with data that were available in year $t + n$; so $CI_{2007, 2021}$ uses the data for the period 2007Q1–2020Q4 that were available in 2021 for example.

The correlation between $OG_{t, t+1}$ and $CI_{t, t+1}$ is *0.90*, suggesting that the two variables evaluate the same object but do not tell exactly the same story. Given the large set of variables from which it is calculated, it is very likely that the CI contains information beyond that in the OG. Assuming that to be the case, the method in this section is based on the notion that the CI can correct the OG and so may predict the direction in which the OG will be revised, and from there the direction of SB revisions.

I define *consistency* between the OG and the CI as the case when both are positive, negative or zero. As a value of exactly zero is very unlikely for either variable and since uncertainty makes it wise to assume that the economy is operating at potential as long as the OG or the CI are in the vicinity of zero, I define zero as the case when $|OG| \leq OGz$ and $|CI| \leq CIz$ for the OG and the CI, where the OGz and COz thresholds for zero are greater than *0* and small. Positive is then defined as above the threshold for zero and negative as below the negative of the threshold.

I define $CI_{t, t+n}$ larger than $OG_{t, t+n}$ as the case when $CI_{t, t+n}$ is positive while the $OG_{t, t+n}$ is zero or negative, or $CI_{t, t+n}$ is zero and $OG_{t, t+n}$ is negative. I define $CI_{t, t+n}$ smaller than $OG_{t, t+n}$ in a similar way.

I set $OGz = 0.75$ and $CIz = 0.15$. In theory, and ignoring the evidence that OG data are not precise and are of dubious reliability, it does not seem acceptable to declare that the economy is at its potential when the output gap is greater than 1pp in absolute terms. Neither does it seem particularly acceptable not to declare that the economy is above or below its potential if the OG is greater or smaller than, say, 0.80pp or -0.80pp, but in practice the imprecision of the data means the threshold for declaring that the economy is above potential cannot be much lower than 0.80, which gives the threshold value. Setting the threshold for the CI is more difficult; while a CI value one standard deviation above its mean should probably indicate overheating in the economy, it is not clear what the minimum level would be where that same conclusion could be drawn. It seems safe to choose 0.15 for this, though it might perhaps be too low.

With these assumptions, it turns out that in the three cases where $CI_{t, t+1}$ is smaller than $OG_{t, t+1}$ in 2012-2014, the OGs were revised downwards as expected in spring 2021, all of them turning from surplus to deficit from spring 2020 to spring 2021. It should be noted, however, that over the successive data vintages, the three $OG_{t, t+1}$ will be revised both upwards and downwards, with the first upward revision already being made one year later in $t + 2$. The latest revisions lead to the expected improvement in the SB, as $SB_{2012, 2013} = 0.17$ and $SB_{2012, 2021} = 0.92$; $SB_{2013, 2014} = -0.39$ and $SB_{2013, 2021} = 0.75$; and $SB_{2014, 2015} = 0.16$ and $SB_{2014, 2021} = 0.84$. In these three cases, $OG_{t, t+1} > 1.25$.

There are two inconsistencies of the type $CI_{t, t+1}$ larger than $OG_{t, t+1}$ in 2011 and 2016. For 2016, $OG_{2016, 2017}$ is only 0.29 and though it was revised upwards as expected and remained above 0.29 for most of the successive data vintages, it was revised downwards into negative territory in spring 2021. The SB is revised downwards as expected but not because of the change in the OG, which moved against expectations in spring 2021. The SB moves from 0.25 in spring 2017 into a deficit of -0.17 in spring 2021 because the 2016 GB, which affects the SB in a one-to-one way, was revised downwards by 0.68pp from 0.27 in spring 2017 to -0.41 in spring 2021. This is further proof that it is especially difficult to rely on the sign of an estimate of the OG, the SB and even the GB when it is fairly close to zero.

$OG_{2011, 2012}$ is fairly low at -1.53. Again counter to expectations, the OG was revised downwards in spring 2021 to -2.33, leading to an upward revision of the SB from -0.20 to 0.72; the GB barely changed. However, with the exception of spring 2021, all the OG estimates are upward revisions of the initial -1.53.

Except for the reference years 2005-2007, all the previous year estimates of the OG by the EC were revised downwards in spring 2021 because of the methodological change discussed earlier that was specific to Estonia, and as stated before there were downward revisions from spring 2020 to spring 2021 for all the years including the first ones.

There is thus some evidence that the CI may help predict the direction of change in the OG and from there, the direction of change in the SB. There are the usual caveats, as a lot of uncertainty is involved, especially when the estimates are close to zero; this exercise is done with a small number of observations; and the time horizons differ for all the reference years, as the direction of revision is determined at the same moment in spring 2021 for all the years, implying that the numbers for more recent years like 2019 might not have had enough time to be revised by very much. Consistency comparisons are also sensitive to the values of *OGz* and *CIz*, and there does not seem to be an obvious and useful way to parameterise these values. Importantly, deep methodological changes in the past and possibly in the future in how the OG is estimated might make it impossible to devise any simple means of predicting the direction in which an OG estimate will ultimately be revised or the sign it will ultimately take.

With these caveats in mind, I suggest a method for developing some trust in the SB estimates by using both the thresholds from the previous section and the OG-CI consistency comparisons in the current section. The method is summarised in Appendix 2.

There are four cases. The SB estimate can be in deficit but above the threshold, in which case a consistency comparison of the OG and the CI may provide further trust for the estimate or may suggest a sign change; or it may be below the threshold, in which case OG-CI consistency may or may not reinforce trust in the estimate; and two similar cases for when the SB estimate is positive. We can apply this method to estimates of the previous year and current year SB and I illustrate it using previous year estimates. The reference years are 2010–2019 because of the availability of CI data (see above).

The previous year SB is negative but above the threshold (2010-2011, 2013, 2017)

The estimates for the previous year SB for 2010-2011, 2013 and 2017 are all negative at -0.40, -0.20, -0.39 and -1.21, but they are not lower than the -1.70 threshold and the rule of thumb suggests we cannot be sure that there is indeed a deficit. So we want to know the likelihood of those negative numbers turning into a surplus in a later estimate.

There is OG-CI consistency for 2010 and 2017, with both indicators declaring the output to be below potential in 2010 and above potential in 2017. This gives us some trust that the SB will continue to be estimated as being in deficit in both years, as OG-CI consistency suggests in general that the OG should remain constant as positive, negative or zero as defined earlier.

However, even if it remains constant in that sense, the OG can be revised enough to change the sign of the SB. If both the OG and the CI are negative and the OG is revised downwards for example, OG-CI consistency remains, the OG remains negative and the OG downward revision might force the SB into surplus.

Four years after the previous year estimate, the 2010 SB was revised to become positive, and it has kept that sign ever since, jumping a whole percentage point from spring 2020 to spring 2021. By contrast, the 2017 SB has kept its negative sign. This might be because not enough time has passed for it to change its sign, though all the estimates are lower than the previous year estimates, which may indicate that the sign will remain negative in future estimates.

Hence there is mixed evidence here for the notion that if the SB is negative but above its threshold and the OG-CI is consistent, then there is some trust the deficit will remain.

For 2011 and 2013 there is OG-CI inconsistency. In 2011, the CI is larger than the OG as the CI is zero and OG is negative, so an improvement in the OG is expected. In that case there is a lot of trust that the SB is indeed in deficit. Successive SB estimates for 2011 keep the negative sign, but it eventually turns and stays positive for a period before becoming negative again and returning back to positive in spring 2021. As $SB_{2011,2012} = -0.20$ is close to zero, this case is not a very strong argument against the usefulness of the OG-CI consistency criterion but it nevertheless contradicts its prediction.

Successive estimates of the OG for 2011 were all above the previous year estimate, as expected, until there was a downward revision of more than 2.30 percentage points from spring 2020 to spring 2021, putting tremendous pressure on the SB to turn positive, which it did.

Once more, the methodological change in spring 2021 makes it very hard to find simple ways of building trust in the sign of the SB estimates. The notion that if SB is negative but above its threshold, and the CI is larger than the OG then there is great trust that the deficit will remain is supported up to spring 2020 but contradicted in spring 2021.

$SB_{2013,2014}$ is - *0.39*, which is also not far from zero, while $OG_{2013,2014}$ is *1.27* and $CI_{2013,2014}$ is *0.11*, meaning they are inconsistent, since the OG suggests the economy was overheating, while the CI says it was running at potential, as the OG is positive and the CI is zero. If we trust the CI more than the OG, we may expect the OG to be revised downwards eventually, leading to an improvement in the structural balance and possibly turning it to surplus. This means that it is likely or very likely that the 2013 SB deficit will become a surplus.

Indeed, after going up from 1.27 to 1.64 in 2015, the 2013 OG fell steadily to 1.01 in spring 2020, making it almost consistent with the CI, until it became negative in spring 2021 at -0.82. Given the assumed *OGz*, this number means that the economy in 2013 was not after all running above its potential but below it. The SB estimates oscillate but remain negative until the SB turns from a very small deficit of -0.15 in spring 2020 to a surplus of 0.75, helped by the drastic change in the OG in spring 2021.

This provides evidence in favour of the rule that if the SB is negative but above its threshold and the CI is smaller than the OG then it is likely or very likely that the deficit will become a surplus. The OG and SB with their ups and downs eventually moved as expected until spring 2020, and it was the methodological revision in spring 2021 that finally turned the SB deficit into a surplus. So there is evidence in favour of the method, but the data revision that confirmed the method happened because of the methodological change.

The previous year SB is negative and below the threshold (2018 and 2019)

The estimates of the previous year SBs for 2018 and 2019 are much lower than -1.70 at -2.17 and -2.40, so the rule of thumb suggests they are very likely to be in deficit.

Both the OG and the CI are consistent for 2018 and 2019, with the OG at 3.33 and 4.28 and the CI at 0.75 and 0.56, indicating an overheating economy and so giving no sign that the OG might become zero or negative. Putting these three pieces of evidence together means we can have a lot of trust that the 2018 and 2019 deficits are indeed deficits.

Checking this prediction shows that in spring 2021 the estimates of the previous year OG for 2018 and 2019 declined by 1.17 and 1.49pp, which are very substantial changes, but both the OGs still remained very large in spring 2021 at 2.16 and 2.79. In the 2021 vintage of the CI, the values for 2018 and 2019

remained almost unchanged. The OG and the CI remain consistent in their story of an overheating economy.

However, downward revisions of the OG can only lead to increases in the SB, and that happened in spring 2021. Such changes, nevertheless, apply when the initial SB levels are very low. Deficits below the threshold are more likely to resist downward revisions of the OG.

The 2018 and 2019 SBs were revised upwards in spring 2021, with a very substantial revision for 2019. The latest 2018 and 2019 estimates are now -1.61 and -1.28, which is above the rule of thumb but still clearly negative. This means that our prediction that the 2018-2019 deficits would remain as deficits still holds when the latest data are used.

This of course could be because there has not been enough time for the OG to be revised downwards sufficiently and the SB to be revised upwards. However, even after a lot more time has passed, it seems very unlikely that those SB deficits will become surpluses.

Assuming that the spring 2021 GB and OO estimates do not change, that the semi-elasticity is not revised and that no methodological changes occur, the spring 2021 estimates of the 2018 and 2019 OGs would have to be revised downwards by 3.30 and 2.63pp respectively in order for the spring 2021 SB deficits to be revised to exactly zero.

Such sharp downward revisions are, however, very unlikely. The average revision up to spring 2021 for the period 2008-2015, which is sufficiently far back that each previous year estimate has been updated at least five times, and for which the estimate of the previous year OG for all the reference years was revised downwards in spring 2021, is -1.18pp. This would not be sufficient to turn the 2018 and 2019 SBs that are currently estimated to be deficits into surpluses in the next five years[12]. Hence the method for developing trust in the SB estimates seems trustable in this case.

There is then evidence in favour of the rule that if the SB is negative and below its threshold and OG-CI is consistent, then there is a lot of trust that the SB deficit will remain a deficit. This evidence is not driven by the methodological change in spring 2021, but is found despite the methodological change, which runs counter to the SB remaining in deficit as it put upward pressure on it through large downward revisions of the OG.

Calculating how much the OG would have to change from the latest estimate so that the SB would change sign while everything else stays the same would provide yet another method for developing trust in the estimates.

The previous year SB is positive but below the threshold (2012, 2014-2016)

The estimates for the previous year SB for 2012 at 0.17 and 2014-2016 at 0.16, 0.58, and 0.25 are all positive but below the rule of thumb of 4.00. To gain some trust for these surpluses, we compare the OGs with the CIs. They are only consistent for 2015, suggesting the economy was at its potential in that year. This consistency provides some trust that the 2015 surplus will remain a surplus. In spring 2021, the SB actually increased from 0.58 to 0.75, though there were two early negative estimates that were very close to zero. The OG oscillated until it became negative, putting upward pressure on the SB. This evidence may favour the notion that if the SB is positive but below its threshold and OG-CI is consistent, then there is some trust that the SB surplus will remain a surplus; however, had the

[12] From the study of revisions, there is really no significant sign that letting more time pass increases the chances that the accumulated revisions of the OG will be larger or smaller.

OG been revised by the same amount of -1.01 but in the opposite direction, $OG_{2015,\,2021}$ would be 1.35 and the $SB_{2015,\,2021}$ would be negative at -0.23.

The CI is smaller than the OG for 2012 and 2014, suggesting that the OG estimate might be revised downwards, improving the SB. We thus have great trust that the 2012 and 2014 SBs are really in surplus.

$SB_{2012,\,2013}$ is *0.17*, which is very similar to $SB_{2014,\,2015} = 0.16$, and both of them are in the vicinity of zero. The 2012 SB turned negative in spring 2014 but was very close to zero and remained negative and never lower than -0.25 until spring 2020, with the exception of spring 2019, when it was 0.05. Then there was a large jump to a value of 0.92 in spring 2021. While the sequence of estimates does not seem to support the method, the latest estimate, being positive, does. The OG for 2012 remained above 0.75% of potential GDP until spring 2021, when it became -0.60 and thus achieved consistency with the CI in declaring that the 2012 OG was zero. This is evidence in favour of the method.

The sequence of the 2014 SB is quite similar, as the SB eventually moved into a deficit extremely close to zero, then returned to a small positive level in spring 2019, and increased sharply in spring 2021. It is true that in the current estimate it is still in surplus, but it is also true it was not so for the whole period, though the negative values were in the two cases very close to zero. The story of the 2014 OG is essentially the same as that of the 2012 OG, as it is always above 0.75% but becomes zero in spring 2021 at -0.04.

The 2012 and 2014 SB estimates give mixed or maybe favourable support to the notion that if the SB is positive but below its threshold and the CI is lower than the OG then there is a lot of trust that the SB surplus will remain a surplus.

The 2012 and 2014 SB examples suggest an extra rule of thumb that may bring realism to the ability to make predictions: when an SB estimate is very close to zero, discretion may be the better part of valour and so predictions should not be attempted as even if the SB estimate does not change by much, it can easily change its sign. There may be no way to gain trust about the sign of such an estimate.

For 2016, the previous year estimate of 0.25 is below the rule of thumb of 4.0 and fairly close to zero, and the CI is positive while the OG is zero, suggesting the OG will go up, pulling the SB down. The algorithm suggests that it is very likely the 2016 SB will prove to be in deficit.

The SB was always in deficit from spring 2018 to spring 2021, even with the effect of the decline in the OG from the penultimate estimate to the latest; 2016 is an exception because the downward revision of the OG from spring 2020 to spring 2021 was not so drastic. The OG went up, as initially expected, from spring 2017 to spring 2019, and then it decreased until spring 2021, becoming negative (actual value) but zero. This actually means there was consistency, because by spring 2021 the CI had become zero.

This case gives a somewhat favourable signal that if the SB is positive but below the threshold and the CI is larger than the OG, then it is very likely the SB is actually in deficit. The SB indeed turned to deficit and remained so but this was in spite of the OG ultimately decreasing, which runs counter to the expectation from the initial consistency comparison.

There are no cases in the database, of the *previous year SB being positive and above the threshold*.

Discussion

The evidence for the suggested method is, overall, mixed or favourable with caveats, which continue to apply for the thresholds. Equally, the latest SB estimate may confirm our sign prediction given the data available one year after the reference year, but it is possible that the SB sign will change from the previous year estimate up to the most recent data vintage and contradict the prediction, even if only temporarily.

It is also possible that the initial sign prediction that is made one year after the reference year will be confirmed by the latest data vintage, but an updated prediction would give a different result that then turns out to be wrong. The prediction for 2016 was that the SB surplus was actually a deficit, and all the estimates after the previous year estimate are indeed negative, but in spring 2019, while the SB was negative as predicted, the OG-CI comparison changed from CI larger than OG, as it was positive and OG was zero to become smaller, as CI was zero and OG positive. As $SB_{2016,\ 2019} = -0.75$ is not that low, and as it was expected in spring 2019 that the OG would decrease, the updated prediction could very well have been that the 2016 SB was likely to become a surplus, which it never did[13].

With these caveats in mind, there is some evidence that the method helps predict the direction of change of the OG and SB estimates and whether an SB surplus will remain a surplus and a deficit will remain a deficit.

The method even helped in some, but not all, of those cases in which the SB of the previous year was very close to zero. $SB_{2013,\ 2014} = -0.39$ is apparently close to zero but the method still helped. $SB_{2016,\ 2017} = 0.247$ is even closer to zero and the method helped too. When the absolute value of $SB_{t,\ t+1}$ is lower than or equal to perhaps 0.20, it might be better to exercise discretion.

Revisions to the government balance may or may not help our predictions come true. In the case of the 2013 deficit, a little of the change in the SB can be attributed to the GB revisions. The GB changes sign from -0.18 in spring 2014 to +0.18 in spring 2021, helping confirm our prediction that the previous year SB deficit would eventually turn into a surplus. The overall change in the GB from spring 2014 to spring 2021 is 0.36pp, which is much smaller in magnitude than the -2.09pp change in the 2013 OG estimates over the same period. We can say that the OG revisions contributed more to our prediction being realised than the GB revisions did.

5. Conclusions

The structural budget balance estimates for Estonia from the EC, OECD and IMF are not precise, and those of the EC are not reliable while the apparent reliability of those by the OECD and the IMF can easily be questioned. As the rules about the structural budget, which is a central concept in the EU fiscal framework, and about the MTO and the pace for moving towards it are all set in tenths of a percentage point of output, those estimates are not a sound basis for setting fiscal policy and for assessing compliance with the fiscal rules. Another main conclusion is that the structural budget balance is a highly inadequate building block for a set of fiscal rules[14]. Furthermore, different

[13] To be coherent with our method, the updated predictions should consider *updated thresholds*, which are calculated from descriptive statistics in which the *initial estimate* is published one or more years after the previous year estimate. For example, if eight years have passed since the previous year estimate, meaning nine years since the reference year, and the time horizon from the initial estimate is one year, then the statistics are based on the differences $SB_{2005,\ 2015} - SB_{2005,\ 2014}$, $SB_{2004,\ 2014} - SB_{2004,\ 2013}$, and so on.

[14] This point has been consistently made by, among others, the European Fiscal Board in its many reports. See https://ec.europa.eu/info/business-economy-euro/economic-and-fiscal-policy-coordination/european-fiscal-board-efb_en#reports-and-publications.

institutions may provide contradictory estimates for whether the sign of the structural balance is positive or negative, leading to contradictory policy recommendations.

No institution is clearly much more precise than the others at estimating the structural balance, and none manages without making large revisions. Since the general level of the estimates of structural balance is much higher for the OECD than for the EC by 0.89pp on average, it is possible that one or even both of the two institutions only produces biased estimates of the structural balance.

Sign revisions are common in the EC data in all revision categories, including revisions of the previous year estimates one year later. The apparent sign reliability of the OECD might be a consequence of general overestimation and that of the IMF might be a consequence of the small sample size. The SB data from the IMF are also larger and further away from zero than those of the EC, and there would be more sign changes if those estimates were closer to zero, unless the IMF data revisions were smaller in size.

It could also be that different institutions are estimating the value of slightly different concepts of the output gap, which is the basis for the calculation of the structural balance. This then shows that the structural balance depends on a conceptually complicated quantity, making it unfit for its current role as the central macroeconomic concept in the European Union fiscal framework and in the specific fiscal rules of Estonia.

From some descriptive statistics with the data at hand, it is possible to suggest rules of thumb for assessing whether a structural deficit or surplus is indeed a deficit or surplus, or for at least providing some confidence that a surplus will not be later revised into a deficit or vice versa. These rules of thumb are subject to the caveats of imprecision and the unreliability of the structural balance estimates, and of short samples. The suggested thresholds for Estonia are -3.40 for the current year deficit estimate, 5.15 for the current year surplus estimate, -1.70 for the previous year deficit estimate, and 4.00 for the previous year surplus estimate.

Respecting the thresholds may give strong confidence that the sign of the SB will not be changed. However, most of these thresholds might be too large in their absolute value to be practically useful.

As the EC estimates of structural balance change sign easily and this often happens more than once for the same reference year over a fairly short sequence of revisions, it is better to look at other macroeconomic variables to give some trust beyond that provided by the thresholds. The heat map composite indicator summarises many macroeconomic variables and it proves somewhat good for providing trust in the estimates.

I propose a simple algorithm for using the heat map that takes the structural balance estimate thresholds and an evaluation of the consistency between the output gap and the composite indicator numbers. This method provides some help in predicting the direction of change in the estimates of the output gap and the structural budget balance and so can give some confidence that a structural deficit or surplus will remain as such in later data releases. When the absolute value of previous year estimates are close to zero, at lower than or equal to perhaps 0.20, it may be better not to predict anything.

Estonia is a very open economy with a very liberal labour market. Macroeconomic quantities can be especially volatile and so can be hard to predict. That might explain the limitations of the method and also of the estimates. The method might work better in less volatile countries, but in those countries there may be less need for it.

The evidence in the paper provides very strong support for the recommendation that structural balance estimates should be excluded from fiscal frameworks. They are a very poor indicator of the state of public finances and are of little use for setting economic policy.

The estimates by the European Commission of the output gap for Estonia changed drastically in spring 2021. That these estimates are so radically sensitive to one single methodological change and that they are normally revised by very considerable amounts shows they are unfit for assessing the state of the economy and for setting economic policy.

The problem of the imprecision and unreliability of estimates is further compounded in Estonia, as its State Budget Act (see footnote 1) sets rules about a *cumulative structural budget balance*[15]. In light of the evidence in this paper, the rule stating that "[t]he status of the cumulative structural budget position is fixed as at the end of the previous budgetary year in the stability programme and is not subsequently adjusted" seems utterly unjustifiable[16]. If it is respected, such a rule could easily lead to a sequence of inappropriate policies based on wrong data.

While they are still considered, estimates of structural budget balance should always be read in the context of a broad and reliable view of the position of the economy in the business cycle, such as the one given by a good set of macroeconomic indicators that can be easily summarised into a composite indicator, and it should be kept in mind how imprecise and unreliable the output gap estimates can be. A similar logic applies to the issue of whether the fiscal stance is pro-cyclical or counter-cyclical.

The work in this paper can be extended to data vintages from the Ministry of Finance and from the central bank, the estimates of which play a central role in policy discussions, and also to the data from other countries and sources such as private banks.

[15] State Budget Act §6 (1^2): "Cumulative structural surplus was generated in case the amount of the structural budget positions of the general government sector created as of 2014 until the year preceding the current budgetary year denominated in euros is positive. If the amount is negative, cumulative structural deficit was generated".
[16] State Budget Act §6 (1^3), https://www.riigiteataja.ee/en/eli/ee/529122017006/consolide/current

References

Darvas, Zsolt (2016): *The structural budget balance limbo*, Bruegel, https://www.bruegel.org/2016/04/the-structural-budget-balance-limbo.

Darvas, Zsolt (2019): *Uncertainty over output gap and structural-balance estimates remains elevated*, Bruegel, https://www.bruegel.org/2019/06/uncertainty-over-output-gap-and-structural-balance-estimates-remains-elevated.

Draghi, Mario and Emmanuel Macron (2021): *The EU's fiscal rules must be reformed*, Financial Times, December 23rd, 2021.

European Commission (2018): *Recent developments in the fiscal surveillance framework*, Report on Public Finances in EMU 2018.

Fiscal Council (April 2018), *Fiscal Council Opinion on the Spring Forecast 2018 of the Ministry of Finance and on the Targets for the Structural Budget Position for 2019-2022*, https://media.voog.com/0000/0036/0984/files/2018_spring_fiscal_council_opinion.pdf.

European Commission (2019): *Vade Mecum on the Stability & Growth Pact*, Luxembourg: Publications Office of the European Union.

European Commission (n.d.): economic and finance glossary at https://ec.europa.eu/info/sites/default/files/economy-finance/glossary.pdf.

European Fiscal Board (2021): *Annual Report 2021*, Brussels.

Mourre et al. (2013): *The cyclically-adjusted budget balance used in the EU fiscal framework: an update*, European Commission, Economic Papers 478.

Mourre et al. (2014): *Adjusting the budget balance for the business cycle: the EU methodology*, European Commission, Economic Papers 536.

Riigikogu (2022): *State Budget Act*, https://www.riigiteataja.ee/en/eli/ee/529122017006/consolide/current.

Vicente, Ricardo (2020): *The Business Cycle, Potential Output and the Output Gap: Different Theories, Different Rules*, Fiscal Council Occasional Paper 1, https://media.voog.com/0000/0036/0984/files/FC_Occasional_Paper_01.pdf.

Appendix 1: Tables

Tables 1 – 3: European Commission Spring Estimates

1. General government budget balance (percentage of GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S 2006	1.60	1.40													
S 2007	2.30	3.80	3.70												
S 2008	1.80	3.40	2.80	0.40											
S 2009	1.50	2.90	2.70	-3.00	-3.00										
S 2010	1.60	2.50	2.60	-2.70	-1.70	-2.40									
S 2011	1.62	2.45	2.54	-2.85	-1.76	0.14	-0.62								
S 2012	1.62	2.46	2.39	-2.95	-2.01	0.26	1.03	-2.35							
S 2013	1.62	2.46	2.39	-2.95	-1.98	0.20	1.17	-0.27	-0.29						
S 2014	1.62	2.46	2.39	-2.95	-1.95	0.19	1.11	-0.24	-0.18	-0.54					
S 2015	1.14	2.93	2.51	-2.66	-2.17	0.18	1.19	-0.22	-0.21	0.62	-0.23				
S 2016	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.77	0.45	-0.12			
S 2017	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.68	0.10	0.27	-0.26		
S 2018	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.68	0.07	-0.29	-0.29	-0.01	
S 2019	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.69	0.07	-0.33	-0.38	-0.55	-0.33
S 2020	1.13	2.89	2.74	-2.64	-2.18	0.19	1.08	-0.28	0.18	0.70	0.14	-0.52	-0.77	-0.56	-0.32
S 2021	1.13	2.89	2.74	-2.64	-2.18	0.19	1.08	-0.28	0.18	0.70	0.14	-0.41	-0.72	-0.56	0.08

2. Output gap (percentage of potential GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S 2006	0.40	1.10													
S 2007	-0.10	1.70	0.70												
S 2008	1.90	4.70	4.00	-0.50											
S 2009	4.40	8.80	10.40	3.30	-8.60										
S 2010	4.30	8.30	11.00	4.50	-10.10	-8.60									
S 2011	7.38	10.73	11.76	3.21	-11.14	-7.66	-3.15								
S 2012	7.02	10.75	12.92	5.52	-9.83	-7.67	-1.53	-1.85							
S 2013	5.47	9.29	12.05	4.64	-9.37	-5.96	0.50	1.37	1.22						
S 2014	5.76	9.76	12.39	4.79	-9.55	-7.04	0.28	2.59	1.27	0.66					
S 2015	6.31	10.73	13.95	5.22	-9.61	-7.04	-0.69	2.19	1.64	1.31	0.99				
S 2016	6.75	11.42	14.83	6.00	-8.99	-6.65	-1.20	2.11	1.41	1.78	0.34	-0.15			
S 2017	6.37	10.96	14.35	5.70	-9.02	-6.58	-0.79	1.83	1.31	1.95	1.02	0.29	0.11		
S 2018	6.27	10.71	14.01	5.38	-9.29	-6.81	-0.99	1.59	1.35	1.87	0.84	0.42	2.08	2.88	
S 2019	6.19	10.61	13.86	5.22	-9.48	-7.04	-1.21	1.31	1.01	1.25	0.34	1.09	2.67	3.33	2.79
S 2020	6.12	9.67	12.43	4.40	-9.58	-6.27	0.03	1.68	1.01	1.39	0.44	0.35	2.63	3.86	4.28
S 2021	4.64	8.28	10.95	2.81	-11.23	-8.33	-2.33	-0.55	-0.82	-0.04	-0.67	-0.28	1.76	2.16	2.79

3. Structural budget balance (percentage of potential GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S 2006	1.50	1.10													
S 2007	2.40	3.30	3.50												
S 2008	1.20	1.00	1.30	0.40											
S 2009	0.20	-0.70	-0.80	-4.10	-1.00										
S 2010		-0.90	-1.10	-4.30	-0.60	-2.10									
S 2011			-1.40	-4.00	-0.20	-0.40	-0.90								
S 2012				-4.80	-0.90	-0.50	-0.20	-0.80							
S 2013	-0.01	-1.18	-1.53	-4.50	-1.08	-1.12	-0.65	0.17	-0.18						
S 2014	-0.10	-1.32	-1.64	-4.55	-0.96	-0.80	-0.52	-0.05	-0.39	-0.47					
S 2015						0.32	-0.04	-0.23	-0.77	0.16	-0.38				
S 2016						0.15	0.18	-0.25	-0.63	0.10	0.58	0.05			
S 2017						0.11	0.00	-0.12	-0.58	-0.07	-0.06	0.25	-0.31		
S 2018						0.22	0.09	-0.01	-0.60	-0.03	-0.02	-0.37	-1.21	-1.29	
S 2019						0.62	0.24	0.05	-0.49	0.20	0.18	-0.75	-1.68	-2.17	-1.69
S 2020						0.28	-0.43	-0.16	-0.15	0.14	0.21	-0.58	-2.05	-2.44	-2.40
S 2021						1.28	0.72	0.92	0.75	0.84	0.75	-0.17	-1.57	-1.61	-1.28

Tables 4 – 6: Consecutive Revisions (source: EC)

Note: Each row corresponds to the difference between two consecutive estimates; for 2006 for example, the row t + 1 is the difference between the spring estimates published in 2007 and 2006, the row t + 2 is the difference between the spring estimates published in 2008 and 2007, and so on.

4. Using the data for the general government budget balance (percentage of GDP):

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
t+1		2.40	-0.90	-3.40	1.30	2.54	1.65	2.08	0.10	1.16	0.68	0.40	-0.03	-0.54	0.01
t+2	0.70	-0.40	-0.10	0.30	-0.06	0.12	0.14	0.03	-0.03	0.14	-0.34	-0.56	-0.10	-0.01	0.40
t+3	-0.50	-0.50	-0.10	-0.15	-0.25	-0.07	-0.06	0.02	0.04	-0.09	-0.03	-0.04	-0.39	0.01	
t+4	-0.30	-0.40	-0.06	-0.10	0.03	-0.01	0.08	-0.04	0.00	0.00	0.00	-0.19	0.05		
t+5	0.10	-0.05	-0.15	0.00	0.03	0.00	-0.03	0.00	0.00	0.01	0.07	0.11			
t+6	0.02	0.01	0.00	0.00	-0.22	0.00	0.00	0.00	0.00	0.02	0.00				
t+7	0.00	0.00	0.00	0.30	-0.02	0.00	0.00	0.00	0.35	0.00					
t+8	0.00	0.00	0.12	-0.01	0.00	0.00	0.00	-0.03	0.00						
t+9	0.00	0.47	0.20	0.00	0.00	0.00	-0.08	0.00							
t+10	-0.48	-0.03	0.00	0.00	0.00	0.00	0.00								
t+11	-0.01	0.00	0.00	0.00	0.01	0.00									
t+12	0.00	0.00	0.00	0.02	0.00										
t+13	0.00	0.00	0.02	0.00											
t+14	0.00	-0.01	0.00												
t+15	0.00	0.00													
t+16	0.00														

5. Using the data for the output gap (percentage of potential GDP):

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
t+1		0.60	3.30	3.80	-1.50	0.94	1.62	3.23	0.06	0.65	-0.65	0.44	1.97	0.45	1.48
t+2	-0.50	3.00	6.40	1.20	-1.04	-0.01	2.03	1.22	0.36	0.46	0.68	0.12	0.60	0.53	-1.49
t+3	2.00	4.10	0.60	-1.29	1.32	1.71	-0.22	-0.40	-0.22	0.18	-0.18	0.67	-0.05	-1.70	
t+4	2.50	-0.50	0.76	2.31	0.46	-1.09	-0.97	-0.08	-0.10	-0.08	-0.50	-0.73	-0.87		
t+5	-0.10	2.43	1.15	-0.88	-0.19	0.00	-0.51	-0.28	0.04	-0.62	0.10	-0.63			
t+6	3.08	0.02	-0.87	0.15	-0.05	0.39	0.41	-0.24	-0.34	0.14	-1.11				
t+7	-0.36	-1.46	0.35	0.43	0.61	0.07	-0.20	-0.27	0.00	-1.43					
t+8	-1.55	0.47	1.55	0.78	-0.03	-0.23	-0.23	0.37	-1.83						
t+9	0.29	0.98	0.89	-0.30	-0.27	-0.23	1.24	-2.24							
t+10	0.56	0.69	-0.48	-0.32	-0.19	0.77	-2.36								
t+11	0.44	-0.46	-0.34	-0.17	-0.10	-2.06									
t+12	-0.38	-0.25	-0.15	-0.82	-1.65										
t+13	-0.11	-0.11	-1.44	-1.59											
t+14	-0.08	-0.94	-1.48												
t+15	-0.07	-1.39													
t+16	-1.48														

6. Using the data for the structural budget balance (percentage of potential GDP):

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
t+1		2.20	-2.20	-4.50	0.40	1.70	0.70	0.97	-0.21	0.62	0.96	0.20	-0.90	-0.88	-0.71
t+2	0.90	-2.30	-2.10	-0.20	0.40	-0.10	-0.45	-0.21	-0.38	-0.06	-0.64	-0.62	-0.48	-0.27	1.12
t+3	-1.20	-1.70	-0.30	0.30	-0.70	-0.62	0.12	-0.18	0.14	-0.17	0.05	-0.38	-0.36	0.83	
t+4	-1.00	-0.20	-0.30	-0.80	-0.18	0.32	0.49	-0.02	0.05	0.04	0.20	0.17	0.47		
t+5				0.30	0.12	1.11	0.22	0.13	-0.02	0.23	0.02	0.41			
t+6				-0.05		-0.17	-0.18	0.11	0.11	-0.05	0.54				
t+7			-0.10			-0.03	0.09	0.06	0.35	0.69					
t+8		-0.14				0.10	0.15	-0.21	0.89						
t+9	-0.09					0.40	-0.67	1.09							
t+10						-0.34	1.15								
t+11						1.00									

Table 7: Output Gap Revision and General Government Budget Balance Revision: Relative Contributions to Structural Budget Balance Revisions (source: EC)

Notes: The table shows the difference between the absolute values of the consecutive revisions of the deflated output gap and the general government budget balance, where the deflated output gap is the output gap multiplied by the semi-elasticity. For 2006 for example, a consecutive revision is the difference between the data published in 2007 and 2006 (row t + 1), the difference between the data published in 2008 and 2007 (row t + 2), and so on. In the table, a positive value means that the output gap revision contributed more than the general government budget balance revision to the revision of the structural budget balance estimate and a negative value means it contributed less.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
t+1		-2.22	0.09	-2.26	-0.85	-2.25	-1.16	-1.11	-0.09	-0.87	-0.39	-0.20	0.84	-0.32	0.71
t+2	-0.55	0.50	1.82	0.06	0.26	-0.12	0.47	0.33	0.13	0.06	-0.04	-0.51	0.19	0.25	0.33
t+3	0.10	0.73	0.08	0.24	0.14	0.45	0.00	0.16	0.05	-0.01	0.05	0.29	-0.36	0.82	
t+4	0.45	-0.25	0.17	0.60	0.11	0.32	0.35	0.00	0.04	0.04	0.24	0.17	0.37		
t+5	-0.07	0.67	0.19	0.26	0.03	0.00	0.19	0.12	0.00	0.30	-0.03	0.20			
t+6	0.91	-0.01	0.26	0.04	-0.19	0.17	0.18	0.11	0.17	0.05	0.54				
t+7	0.11	0.44	0.10	-0.11	0.26	0.03	0.09	0.13	-0.35	0.69					
t+8	0.47	0.14	0.56	0.34	0.01	0.10	0.11	0.15	0.89						
t+9	0.09	-0.04	0.19	0.13	0.12	0.11	0.52	1.09							
t+10	-0.23	0.28	0.21	0.14	0.09	0.37	1.15								
t+11	0.18	0.20	0.15	0.08	0.04	1.00									
t+12	0.17	0.11	0.07	0.37	0.80										
t+13	0.05	0.05	0.67	0.77											
t+14	0.04	0.44	0.72												
t+15	0.03	0.67													
t+16	0.72														

Tables 8 – 10: Descriptive Statistics of Absolute Consecutive Revisions for Each Reference Year (source: EC)

Notes: The upper panels of the tables show descriptive statistics for the absolute values of consecutive revisions for each reference year, meaning across data vintages but excluding revisions to the current year estimate, which give the difference between the data for year t published in that same year and one year later. In the lower panels, a zero revision is an absolute consecutive revision that is equal to or smaller than 0.01pp; for 2006 for example, years until first zero revision are equal to 1 if there was a zero revision from spring 2006 to spring 2007, equal to 2 if there was a zero revision from spring 2007 to spring 2008, and so on. A value of 1 never arises, as estimates for a given year calculated in the spring of that same year will naturally change one year later. A value of 2 might imply that the estimate for, say, the year 2006 calculated in 2007 was very precise; however, after a zero revision there may be non-zero data revisions, some of which may be large. For each reference year, the average after the first zero revision calculates the mean of the absolute consecutive revisions over all the vintages after the first zero revision; other zero revisions may be included in the average. The lower panels also show the number of observations and the maximum absolute consecutive revision after the first zero revision, if any. Table 10 has no lower panel because there are no structural budget balance zero revisions.

8. Using the data for the general government budget balance (percentage of GDP):

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
N	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.05	0.01	0.40
Average	0.14	0.13	0.06	0.07	0.05	0.02	0.04	0.01	0.06	0.04	0.09	0.22	0.18	0.01	0.40
S.D.	0.23	0.20	0.07	0.11	0.09	0.04	0.05	0.02	0.12	0.05	0.13	0.20	0.15	0.00	0.00
Max	0.70	0.50	0.20	0.30	0.25	0.12	0.14	0.04	0.35	0.14	0.34	0.56	0.39	0.01	0.40
Years until first zero	7	7	6	5	8	4	6	5	4	4	4	-	-	2	1
Average after first zero	0.06	0.06	0.04	0.04	0.00	0.00	0.02	0.01	0.09	0.01	0.04	-	-	0.01	0.40
N after first zero	9	8	8	8	4	7	4	4	4	3	2	-	-	1	1
Max after first zero	0.48	0.47	0.20	0.30	0.01	0.00	0.08	0.03	0.35	0.02	0.07	-	-	0.01	0.40

9. Using the data for the output gap (percentage of potential GDP):

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
N	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
Min	0.07	0.02	0.15	0.15	0.03	0.00	0.20	0.08	0.00	0.08	0.10	0.12	0.05	0.53	1.49
Average	0.90	1.20	1.27	0.85	0.54	0.65	0.91	0.64	0.41	0.49	0.51	0.54	0.50	1.12	1.49
S.D.	0.94	1.16	1.55	0.63	0.53	0.70	0.77	0.68	0.59	0.46	0.36	0.24	0.34	0.58	0.00
Max	3.08	4.10	6.40	2.31	1.65	2.06	2.36	2.24	1.83	1.43	1.11	0.73	0.87	1.70	1.49
Years until first zero	-	-	-	-	-	2	No	-	7	-	-	-	-	-	-
Average after first zero	-	-	-	-	-	0.73	-	-	1.83	-	-	-	-	-	-
N after first zero	-	-	-	-	-	9	-	-	1	-	-	-	-	-	-
Max after first zero	-	-	-	-	-	2.06	-	-	1.83	-	-	-	-	-	-

10. Using the data for the structural budget balance (percentage of potential GDP):

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
N	4	4	4	5	4	10	9	8	7	6	5	4	3	2	1
Min	0.09	0.14	0.10	0.05	0.12	0.03	0.09	0.02	0.02	0.04	0.02	0.17	0.36	0.27	1.12
Average	0.80	1.08	0.70	0.33	0.35	0.42	0.39	0.25	0.28	0.21	0.29	0.40	0.44	0.55	1.12
S.D.	0.42	0.94	0.81	0.25	0.23	0.36	0.33	0.32	0.28	0.23	0.25	0.16	0.05	0.28	0.00
Max	1.20	2.30	2.10	0.80	0.70	1.11	1.15	1.09	0.89	0.69	0.64	0.62	0.48	0.83	1.12

Tables 11 – 13: Descriptive Statistics of Absolute Consecutive Revisions For Each Time Horizon (source: EC)

Notes: The tables show descriptive statistics for each time horizon, meaning across the reference years using the absolute values of consecutive revisions. Row t + 1 refers to revisions between the reference year and one year later; row t + 2 refers to revisions between 1 and 2 years after the reference year, and so on. Then, for each row, different data vintages are used. From the first to the last row, a decreasing average, decreasing S.D. and decreasing maximum may be interpreted as signs that the underlying estimates have become more precise vintage after vintage, with the caveat that the statistics in the last rows are calculated from only a few observations.

11. Using the data for the general government budget balance (percentage of GDP):

	N	Min	Average	S.D.	Max
t+1	14	0.01	1.23	1.02	3.40
t+2	15	0.01	0.23	0.20	0.70
t+3	14	0.01	0.16	0.17	0.50
t+4	13	0.00	0.10	0.12	0.40
t+5	12	0.00	0.05	0.05	0.15
t+6	11	0.00	0.02	0.06	0.22
t+7	10	0.00	0.07	0.13	0.35
t+8	9	0.00	0.02	0.04	0.12
t+9	8	0.00	0.10	0.16	0.47
t+10	7	0.00	0.07	0.17	0.48
t+11	6	0.00	0.00	0.01	0.01
t+12	5	0.00	0.00	0.01	0.02
t+13	4	0.00	0.01	0.01	0.02
t+14	3	0.00	0.00	0.01	0.01
t+15	2	0.00	0.00	0.00	0.00
t+16	1	0.00	0.00	0.00	0.00

12. Using the data for the output gap (percentage of potential GDP):

	N	Min	Average	S.D.	Max
t+1	14	0.06	1.48	1.15	3.80
t+2	15	0.01	1.31	1.56	6.40
t+3	14	0.05	1.05	1.06	4.10
t+4	13	0.08	0.84	0.74	2.50
t+5	12	0.00	0.58	0.66	2.43
t+6	11	0.02	0.62	0.84	3.08
t+7	10	0.00	0.52	0.49	1.46
t+8	9	0.03	0.78	0.64	1.83
t+9	8	0.23	0.80	0.65	2.24
t+10	7	0.19	0.77	0.68	2.36
t+11	6	0.10	0.59	0.67	2.06
t+12	5	0.15	0.65	0.55	1.65
t+13	4	0.11	0.81	0.71	1.59
t+14	3	0.08	0.83	0.58	1.48
t+15	2	0.07	0.73	0.66	1.39
t+16	1	1.48	1.48	0.00	1.48

13. Using the data for the structural budget balance (percentage of potential GDP):

	N	Min	Average	S.D.	Max
t+1	14	0.20	1.23	1.10	4.50
t+2	15	0.06	0.68	0.66	2.30
t+3	14	0.05	0.50	0.45	1.70
t+4	13	0.02	0.33	0.29	1.00
t+5	9	0.02	0.28	0.32	1.11
t+6	7	0.05	0.17	0.16	0.54
t+7	6	0.03	0.22	0.24	0.69
t+8	5	0.10	0.30	0.30	0.89
t+9	4	0.09	0.56	0.37	1.09
t+10	2	0.34	0.74	0.40	1.15
t+11	1	1.00	1.00	0.00	1.00

Tables 14 – 16: OECD Spring Estimates

14. General government budget balance (percentage of GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S2009	1.50	2.91	2.70	-2.97	-5.57										
S2010	1.62	2.50	2.62	-2.76	-1.72	-2.60									
S2011	1.62	2.45	2.54	-2.85	-1.76	0.14	-0.48								
S2012	1.62	2.46	2.39	-2.95	-2.01	0.26	1.03	-1.96							
S2013	1.62	2.46	2.39	-2.95	-1.98	0.20	1.18	-0.27	0.02						
S2014	1.62	2.46	2.39	-2.95	-1.95	0.19	1.11	-0.24	-0.18	-0.24					
S2015	1.14	2.93	2.51	-2.66	-2.17	0.18	1.19	-0.22	-0.21	0.62	0.42				
S2016	1.13	2.91	2.73	-2.66	-2.18	0.19	1.16	-0.26	-0.17	0.77	0.44	-0.39			
S2017	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.68	0.10	0.27	-0.36		
S2018	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.68	0.07	-0.29	-0.29	0.43	
S2019	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.69	0.07	-0.33	-0.38	-0.55	-0.23
S2020	1.13	2.89	2.74	-2.65	-2.18	0.19	1.08	-0.28	0.18	0.70	0.14	-0.52	-0.77	-0.56	-0.32
S2021	1.13	2.89	2.74	-2.65	-2.18	0.19	1.08	-0.28	0.18	0.70	0.14	-0.41	-0.72	-0.56	0.08

15. Output gap (percentage of potential GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S2011	5.02	11.58	15.93	7.19	-9.32	-7.54	-3.80								
S2012	4.32	10.00	13.62	6.14	-10.00	-8.36	-2.39	-2.36							
S2013	4.45	10.01	13.44	5.13	-11.01	-8.86	-2.63	-1.67	-2.91						
S2014	4.53	10.15	13.66	5.41	-10.74	-9.20	-1.90	-0.23	-1.94	-3.56					
S2015	4.72	10.92	15.18	5.75	-11.14	-9.50	-3.60	-1.39	-2.17	-2.48	-2.94				
S2016	4.39	10.70	14.36	5.30	-11.14	-10.12	-4.91	-2.12	-2.50	-1.56	-2.15	-2.14			
S2017	2.75	8.46	12.10	4.19	-10.81	-9.15	-3.23	-0.87	-1.33	-0.77	-1.39	-1.80	-1.10		
S2018	2.48	8.25	12.00	4.19	-10.74	-9.08	-3.25	-0.97	-1.23	-0.79	-1.38	-1.60	0.30	0.93	
S2019	2.58	8.30	11.97	4.15	-10.76	-9.09	-3.22	-0.88	-1.09	-0.71	-1.09	-0.25	1.42	2.18	2.17
S2020															
S2021	5.90	10.31	12.76	4.07	-11.72	-10.16	-4.69	-3.10	-3.52	-2.78	-2.62	-1.86	0.41	1.45	2.74

16. Structural budget balance (percentage of potential GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S2012	0.17	-0.58	-1.15	-3.90	-1.77	-0.58	-0.08	0.00							
S2013	0.26	-0.39	-0.85	-3.44	-1.15	-0.36	-0.06	1.31	0.95						
S2014	0.14	-0.55	-1.03	-3.63	-1.27	-0.26	-0.29	0.96	0.61	1.10					
S2015	-0.22	-0.63	-1.51	-3.47	-1.05	-0.17	0.10	1.05	0.25	1.85	1.56				
S2016	-0.70	-1.91	-3.05	-3.81	1.15	1.79	1.39	1.62	0.79	1.84	1.98	1.16			
S2017	-0.02	-1.02	-2.14	-3.32	0.95	1.29	0.61	1.07	0.26	1.40	1.29	1.45	0.54		
S2018	0.09	-0.94	-2.10	-3.32	0.91	1.25	0.62	1.11	0.21	1.41	1.25	0.79	-0.06	0.12	
S2019	0.05	-0.95	-2.09	-3.30	0.92	1.25	0.60	1.07	0.15	1.37	1.09	0.07	-0.70	-1.47	-1.15
S2020															
S2021	-1.29	-1.75	-2.32	-3.20	1.47	1.80	1.21	2.02	1.61	2.32	1.88	0.76	-0.56	-1.18	-1.22

Tables 17 – 19: IMF Spring Estimates

17. General government budget balance (percentage of GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S 2011	1.61	3.22	2.86	-2.32	-2.06	0.20	-1.01								
S 2012	1.61	3.22	2.82	-2.30	-2.06	0.40	0.99	-2.09							
S 2013	1.61	3.22	2.82	-2.31	-2.07	0.40	1.70	-0.25	0.41						
S 2014	1.62	2.46	2.39	-2.95	-1.95	0.20	1.21	-0.21	-0.36	-0.42					
S 2015	1.61	2.44	2.36	-2.90	-1.93	0.19	1.05	-0.28	-0.47	0.37	-0.53				
S 2016	1.61	2.44	2.36	-2.90	-1.93	0.19	1.03	-0.41	-0.30	0.75	0.53	0.54			
S 2017	1.61	2.44	2.36	-2.90	-1.93	0.15	1.12	-0.26	-0.17	0.68	0.10	0.27	0.26		
S 2018	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.68	0.07	-0.29	-0.13	-0.35	
S 2019	1.12	2.91	2.72	-2.67	-2.18	0.19	1.16	-0.26	-0.17	0.67	0.07	-0.28	-0.28	0.23	0.24
S 2020	1.12	2.90	2.69	-2.65	-2.17	0.19	1.15	-0.26	-0.17	0.68	0.07	-0.33	-0.38	-0.46	-0.40
S 2021	1.12	2.90	2.69	-2.65	-2.17	0.19	1.15	-0.26	-0.17	0.68	0.07	-0.33	-0.38	-0.46	0.04

18. Output gap (percentage of potential GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S 2012	2.50	7.20	10.87	4.95	-9.80	-7.62	-1.58	-1.38							
S 2013	0.69	5.92	11.69	6.25	-9.36	-7.57	-1.65	-0.70	-0.93						
S 2014	4.13	10.62	11.69	6.25	-9.06	-8.54	-1.83	-0.57	-1.70	-1.16					
S 2015	6.29	12.62	17.52	8.36	-9.58	-9.15	-3.36	-0.62	-0.85	-0.94	-0.63				
S 2016	4.24	10.37	15.09	6.13	-11.29	-10.66	-5.24	-2.09	-2.31	-0.66	-1.30	-1.13			
S 2017	4.10	10.21	14.91	5.95	-11.44	-10.73	-4.55	-2.10	-2.38	-0.81	-0.91	-1.09	-0.50		
S 2018	4.29	10.41	15.13	6.15	-11.27	-10.74	-4.55	-2.10	-1.97	-0.73	-0.65	-0.79	1.27	1.69	
S 2019	4.36	8.90	11.92	3.39	-11.34	-8.90	-3.23	-1.25	-1.73	-1.38	-1.86	-0.97	0.63	1.32	0.95
S 2020															
S 2021	5.14	9.09	11.89	3.73	-10.69	-7.75	-2.02	-0.97	-1.83	-1.19	-1.59	-0.89	1.34	2.50	3.19

19. Structural budget balance (percentage of potential GDP)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S 2016	0.33	-0.69	-2.21	-4.96	1.19	1.76	1.24	1.17	0.73	1.17	1.44	0.79			
S 2017	0.37	-0.64	-2.15	-4.91	1.24	1.75	1.11	1.34	0.89	1.14	0.91	0.93	0.79		
S 2018	-0.16	-0.27	-1.92	-4.76	0.95	1.80	1.15	1.34	0.75	1.12	0.79	0.26	-0.23	-0.81	
S 2019	-0.18	0.19	-0.92	-3.89	0.98	1.15	0.72	1.08	0.68	1.31	1.19	0.32	-0.15	-0.09	
S 2020															
S 2021	-0.41	0.14	-0.90	-3.96	0.74	0.73	0.32	0.99	0.71	1.26	1.10	0.24	-0.48	-1.17	-0.94

Tables 20-21: Descriptive Statistics for the Structural Budget Balance Absolute Consecutive Revisions for Each Reference Year

Notes: The upper panels of the tables show descriptive statistics for the absolute values of consecutive revisions for each reference year, meaning across data vintages but excluding revisions to the current year estimate, which give the difference between the data for year t published in that same year and one year later. In the lower panels, a zero revision is an absolute consecutive revision that is equal to or smaller than 0.01pp; for 2006 for example, years until first zero revision are equal to 1 if there was a zero revision from spring 2006 to spring 2007, equal to 2 if there was a zero revision from spring 2007 to spring 2008, and so on. A value of 1 never arises, as estimates for a given year calculated in the spring of that same year will naturally change one year later. A value of 2 might imply that the estimate for, say, the year 2006 calculated in 2007 was very precise; however, after a zero revision there may be non-zero data revisions, some of which may be large. For each reference year, the average after the first zero revision calculates the mean of the absolute consecutive revisions over all the vintages after the first zero revision; other zero revisions may be included in the average. The lower panels also show the number of observations and the maximum absolute consecutive revision after the first zero revision, if any.

20. Source: OECD

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
N	7	7	7	7	7	7	7	6	5	4	3	2	1
Min	0.04	0.02	0.01	0.00	0.01	0.00	0.01	0.04	0.05	0.01	0.04	0.66	0.64
Average	0.27	0.39	0.50	0.24	0.48	0.42	0.39	0.28	0.31	0.12	0.29	0.69	0.64
S.D.	0.22	0.46	0.51	0.18	0.72	0.65	0.45	0.23	0.22	0.18	0.28	0.03	0.00
Max	0.68	1.29	1.54	0.49	2.20	1.96	1.29	0.57	0.54	0.44	0.69	0.72	0.64
Years until first zero	-	-	-	10	10	9	7	-	-	4	-	-	-
Average after first zero	-	-	-	0.02	-	-	0.01	-	-	0.04	-	-	-
N after first zero	-	-	-	1	0	0	1	-	-	1	-	-	-
Max after first zero	-	-	-	0.02	-	-	0.01	-	-	0.04	-	-	-

21. Source: IMF

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
N	3	3	3	3	3	3	3	3	3	3	3	2	1
Min	0.02	0.05	0.06	0.05	0.02	0.01	0.04	0.00	0.08	0.02	0.12	0.05	0.08
Average	0.20	0.29	0.43	0.36	0.12	0.24	0.21	0.14	0.12	0.08	0.35	0.36	0.08
S.D.	0.23	0.18	0.41	0.36	0.12	0.29	0.17	0.11	0.03	0.08	0.17	0.31	0.00
Max	0.53	0.46	1.00	0.87	0.29	0.66	0.44	0.26	0.16	0.20	0.53	0.67	0.08
Years until first zero	-	-	-	-	-	-	-	6	-	-	-	-	-
Average after first zero	-	-	-	-	-	-	-	0.26	-	-	-	-	-
N after first zero	-	-	-	-	-	-	-	1	-	-	-	-	-
Max after first zero	-	-	-	-	-	-	-	0.26	-	-	-	-	-

Tables 22-23: Descriptive Statistics for the Structural Budget Balance Absolute Consecutive Revisions For Each Time Horizon

Notes: The tables show descriptive statistics for each time horizon, meaning across the reference years using the absolute values of consecutive revisions. Row t + 1 refers to revisions between the reference year and one year later; row t + 2 refers to revisions between 1 and 2 years after the reference year, and so on. Then, for each row, different data vintages are used. From the first to the last row, a decreasing average, decreasing S.D. and decreasing maximum may be interpreted as signs that the underlying estimates have become more precise vintage after vintage, with the caveat that the statistics in the last rows are calculated from only a few observations.

22. Source: OECD

	N	Min	Average	S.D.	Max
t+1	7	0.29	0.76	0.47	1.59
t+2	7	0.01	0.39	0.27	0.69
t+3	7	0.04	0.33	0.23	0.72
t+4	7	0.01	0.34	0.23	0.62
t+5	7	0.04	0.37	0.42	1.29
t+6	7	0.04	0.51	0.64	1.96
t+7	7	0.01	0.47	0.72	2.20
t+8	7	0.01	0.19	0.16	0.48
t+9	6	0.00	0.38	0.54	1.54
t+10	5	0.00	0.51	0.51	1.29
t+11	4	0.02	0.36	0.36	0.90
t+12	3	0.01	0.26	0.30	0.68
t+13	2	0.02	0.06	0.05	0.11
t+14	1	0.04	0.04	0.00	0.04

23. Source: IMF

	N	Min	Average	S.D.	Max
t+1	3	0.14	0.63	0.36	1.01
t+2	3	0.08	0.43	0.25	0.67
t+3	3	0.02	0.06	0.04	0.12
t+4	3	0.03	0.19	0.15	0.40
t+5	3	0.13	0.17	0.03	0.20
t+6	3	0.00	0.07	0.05	0.13
t+7	3	0.01	0.11	0.11	0.26
t+8	3	0.05	0.18	0.18	0.44
t+9	3	0.05	0.33	0.25	0.66
t+10	3	0.02	0.08	0.05	0.15
t+11	3	0.05	0.39	0.35	0.87
t+12	3	0.04	0.47	0.40	1.00
t+13	2	0.46	0.50	0.03	0.53
t+14	1	0.02	0.02	0.00	0.02

Tables 24 – 26: Comparison of Sources

Notes: The minimum, average and maximum are based on the absolute value of revisions; all numbers are percentage point changes. Previous level revision is the difference between year t estimates published in t + 2 and t + 1; current level revision is the difference between year t estimates published in t + 1 and t; previous change revision is the difference between the estimates for the change from t - 1 to t published in t + 2 and t + 1; current change revision is the difference between the estimates for the change from t - 1 to t published in t + 1 and t. These definitions follow Darvas (2016, 2019). Revision to negative or positive is the number of times the sign of an estimate was revised to negative or positive during the reference period.

Sources: EC spring estimates; OECD May or June estimates; IMF April estimates; own calculations.

24. Using the data of the general government budget balance (percentage of GDP):

	Previous level revision			Current level revision			Previous change revision			Current change revision		
	EC	OECD	IMF	EC	OECD	IMF	EC	OECD	IMF	EC	OECD	IMF
	2005-19	2008-19	2010-19	2006-19	2009-19	2011-19	2005-19	2008-19	2010-19	2006-19	2009-19	2011-19
N	15	12	10	14	11	9	15	12	10	14	11	9
Min	0.01	0.01	0.03	0.01	0.02	0.27	0.05	0.05	0.04	0.02	0.08	0.06
Average	0.23	0.18	0.37	1.23	1.15	0.93	0.24	0.23	0.38	1.15	1.20	0.72
Max	0.70	0.56	0.70	3.40	3.85	2.00	0.53	0.53	0.70	3.30	3.64	1.81
Rev. to negative	1	1	2	1	2	3	2	2	2	1	1	1
Rev. to positive	1	1	1	4	4	4	1	1	0	8	6	4

25. Using the data of the output gap (percentage of potential GDP):

	Previous level revision			Current level revision			Previous change revision			Current change revision		
	EC	OECD	IMF	EC	OECD	IMF	EC	OECD	IMF	EC	OECD	IMF
	2005-19	2010-17	2011-17	2006-19	2011-18	2012-18	2005-19	2010-17	2011-17	2006-19	2011-18	2012-18
N	15	8	7	14	8	7	15	8	7	14	8	7
Min	0.01	0.21	0.07	0.06	0.33	0.04	0.07	0.02	0.05	0.15	0.12	0.28
Average	1.31	0.72	0.38	1.48	0.99	0.65	0.70	0.46	0.59	1.23	0.85	0.76
Max	6.40	1.44	0.86	3.80	1.41	1.77	2.30	1.24	1.74	2.70	2.23	1.47
Rev. to negative	1	0	0	0	0	0	0	0	0	0	1	3
Rev. to positive	1	0	0	3	1	1	2	1	1	3	0	0

26. Using the data of the structural budget balance (percentage of potential GDP):

	Previous level revision			Current level revision			Previous change revision			Current change revision		
	EC	OECD	IMF	EC	OECD	IMF	EC	OECD	IMF	EC	OECD	IMF
	2005-19	2011-17	2015-17	2006-19	2012-18	2016-18	2005-19	2011-17	2015-17	2006-19	2012-18	2016-18
N	15	7	3	14	7	3	15	7	3	14	7	3
Min	0.06	0.01	0.08	0.20	0.29	0.14	0.09	0.08	0.03	0.01	0.01	0.35
Average	0.68	0.39	0.43	1.23	0.76	0.63	0.33	0.33	0.36	0.85	0.69	0.55
Max	2.30	0.69	0.67	4.50	1.59	1.01	1.10	0.62	0.55	2.40	1.30	0.67
Rev. to negative	4	0	0	2	2	1	3	2	2	0	1	0
Rev. to positive	0	0	0	2	1	0	1	0	0	6	2	2

Tables 27 – 30: Descriptive Statistics for the Structural Budget Balance Revisions in Later Years Conditional on a Positive or Negative Initial Estimate (source: EC)

Notes: Current year estimate revision n years later is the difference between year t estimates published in t + n and t; the estimate published in year t is the initial estimate. Previous year estimate revision n years later is the difference between year t estimates published in t + 1 + n and t + 1; in this case, the estimate published in year t + 1 is the initial estimate. Each table presents statistics conditional on the initial estimate being positive or negative. Note that the underlying data are the revisions themselves, not their absolute values. All numbers are percentage point changes. Sources: EC spring estimates; own calculations.

27. Conditional on negative current year estimate:

n	Obs.	Mean	S.D.	Mean - S.D.	Mean + S.D.
1	10	0.27	0.85	-0.58	1.12
2	10	0.16	0.88	-0.72	1.04
3	9	0.03	0.75	-0.72	0.78
4	9	0.53	1.07	-0.54	1.61
5	8	0.86	0.86	0.00	1.72
6	6	0.89	0.76	0.14	1.65
7	5	1.08	0.70	0.37	1.78
8	4	1.26	0.64	0.62	1.89
9	3	0.35	0.12	0.23	0.46
10	2	2.00	0.38	1.62	2.38
11	1	3.38	0.00	3.38	3.38

28. Conditional on positive current year estimate:

n	Obs.	Mean	S.D.	Mean - S.D.	Mean + S.D.
1	4	-1.08	2.52	-3.59	1.44
2	4	-2.38	2.13	-4.51	-0.25
3	4	-2.90	1.64	-4.54	-1.26
4	4	-3.18	1.93	-5.11	-1.25
5	2	-2.56	2.34	-4.90	-0.22
6	2	-4.99	0.04	-5.03	-4.95
7	2	-3.71	1.43	-5.14	-2.28
8	1	-2.42	0.00	-2.42	-2.42

29. Conditional on negative previous year estimate:

n	Obs.	Mean	S.D.	Mean - S.D.	Mean + S.D.
1	8	-0.04	0.51	-0.56	0.47
2	7	-0.25	0.44	-0.69	0.19
3	7	0.22	1.37	-1.15	1.59
4	6	0.48	1.09	-0.61	1.57
5	4	0.05	0.37	-0.32	0.42
6	3	0.35	0.12	0.23	0.47
7	3	0.73	0.29	0.44	1.03
8	2	0.39	0.62	-0.23	1.02
9	2	0.80	0.12	0.68	0.92
10	1	1.68	0.00	1.68	1.68

30. Conditional on positive previous year estimate:

n	Obs.	Mean	S.D.	Mean - S.D.	Mean + S.D.
1	7	-0.72	1.05	-1.77	0.33
2	7	-1.27	1.31	-2.59	0.04
3	7	-1.43	1.38	-2.81	-0.06
4	4	-0.26	0.18	-0.44	-0.08
5	4	-0.72	1.23	-1.94	0.51
6	4	-1.71	2.09	-3.80	0.37
7	3	-2.15	1.81	-3.97	-0.34
8	2	-0.42	1.18	-1.60	0.76

Table 31: Composite Indicator Data

Notes: The composite indicator (CI) is the simple average of the standardised values of core inflation, the average wage, the unemployment rate, the employment rate, vacancies, capacity utilisation, demand in construction, demand in industry, demand in services, and economic sentiment; the standardised values are based on averages calculated over quarterly data starting at 2007Q1 and ending at 2021Q1 (Fiscal Council, 2018). Each row corresponds to a different vintage; for example, the 2010 vintage is calculated with data for the period 2007Q1-2009Q4 that was available in 2010. Annual CI is the average of the quarterly CIs of the same year, keeping the vintage the same. A value of, say, 1 or -1 means that the CI is 1 standard deviation above or below its average on a given year, indicating that the economy is above or below its potential.

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
S 2010	0.94	0.24	-1.18										
S 2011	1.18	0.42	-0.96	-0.65									
S 2012	1.30	0.47	-1.04	-0.71	-0.02								
S 2013	1.41	0.51	-1.15	-0.78	-0.03	0.04							
S 2014	1.50	0.52	-1.25	-0.85	-0.05	0.02	0.11						
S 2015	1.56	0.53	-1.32	-0.90	-0.06	0.01	0.09	0.10					
S 2016	1.61	0.53	-1.38	-0.92	-0.06	0.00	0.08	0.09	0.05				
S 2017	1.65	0.52	-1.46	-0.97	-0.09	-0.03	0.05	0.06	0.01	0.25			
S 2018	1.62	0.47	-1.54	-1.04	-0.15	-0.09	-0.01	-0.01	-0.06	0.18	0.65		
S 2019	1.57	0.42	-1.61	-1.10	-0.21	-0.16	-0.08	-0.08	-0.14	0.10	0.57	0.72	
S 2020	1.56	0.39	-1.69	-1.15	-0.25	-0.21	-0.13	-0.13	-0.19	0.05	0.52	0.67	0.56
S 2021	1.60	0.42	-1.68	-1.15	-0.24	-0.19	-0.11	-0.11	-0.17	0.08	0.55	0.71	0.59

Appendix 2: Trusting the SB estimate signs: a simple algorithm

Structural balance confidence thresholds for deficits (surpluses):

- choose to work with the previous year or current year estimates of structural budget balance;
- select those that are in deficit (surplus);
- calculate the mean and standard deviation of the revisions one year ahead, two years ahead, three years ahead, and so on;
- consider constraining the number of horizons by choosing only those for which there is a minimum number of, say, five observations;
- select the highest (lowest) of mean plus (minus) standard deviation, and change its sign;
- this gives the threshold. If a deficit (surplus) is below (above) the threshold for deficits (surpluses) then there is some confidence that the estimate will not change sign in some later revision.

Definitions:

OG is the output gap and *CI* is the composite indicator, which is a simple average of standardised macroeconomic indicators used to generate a broad view of the position of the economy in the business cycle (see section 4.2).

Zero is the case where $|OG| \leq OGz$ and $|CI| \leq CIz$ for the *OG* and the *CI* respectively, with *OGz* and *CIz* > 0 and small; *positive* is when $OG > OGz$ and $CI > CIz$ respectively; and *negative* is when $OG < -OGz$ and $CI < -CIz$ respectively.

OGz should be chosen such that when $OG > OGz$, $OG < -OGz$ or $|OG| \leq OGz$ then it is very credible to say that the economy is above, below or at potential. *CIz* is chosen in a similar way.

Suggested values for Estonia are $OGz = 0.75\%$ of potential GDP and $CIz = 0.15$.

CI and OG being consistent means they are both positive, both negative or both zero.

$CI > OG$ means that the *CI* is positive while the *OG* is zero or negative, or the *CI* is zero and the *OG* is negative.

$CI < OG$ means that the OG is positive while the CI is zero or negative, or the OG is zero and the CI is negative.

Trust tables:

First choose the upper panel for the case SB < 0 and the lower panel for the case SB > 0; then choose the relevant column; finally choose the relevant row.

SB < 0	and below the threshold and…	but above the threshold and…
…CI and OG are consistent then:	great trust that there is a deficit.	some trust that there is a deficit.
…CI > OG: OG is likely to be revised upwards, so SB is likely to be revised downwards then:	greatest trust that there is a deficit.	great trust that there is a deficit.
…CI < OG: OG is likely to be revised downwards, so SB is likely to be revised upwards then:	some trust that there is a deficit.	very likely that there is a surplus, unless it is very close to 0.00

SB > 0	and above the threshold and…	but below the threshold and…
…CI and OG are consistent then:	great trust that there is a surplus.	some trust that there is a surplus.
…CI < OG: OG is likely to be revised downwards, so SB is likely to be revised upwards then:	greatest trust that there is a surplus.	great trust that there is a surplus.
…CI > OG: OG is likely to be revised upwards, so SB is likely to be revised downwards then:	some trust that there is a surplus.	very likely that there is a deficit.

Important note:

The trust tables most probably work better the further away the estimate is from zero. For SBs very close to zero, it is probably wise to expect a lot of sign changes over a lot of years. A suggested value for *very close to zero* for Estonia is 0.20 meaning that when the absolute value of SB is lower than 0.20 it should be considered very close to zero (see paper).

An extra criterion for trust and more:

The sign of the SB may still not be certain. It can be calculated how much the latest OG estimate would need to be revised by so that the latest SB estimate would change sign if the latest estimates of the GB (general government budget balance) and OO (one-off and other temporary budgetary measures) are used, and assuming that they are not revised any more, and assuming as well that the semi-elasticity does not change and there are no methodological changes. Is the magnitude of the necessary OG revision plausible, given that successive OG revisions can be large even many years after the reference year? Successive revisions of the OG 10 years after the reference year in absolute terms were on average 0.77pp with a standard deviation of 0.68 and a maximum of a staggering 2.36pp for example.

Furthermore, new CI vintages are richer in information, and as new vintages are calculated, new OG-CI comparisons may be made to update the expectation about whether the OG might go up or down.